Valuation of Travel Time Reliability in Passenger Transport

Tseng, Y.

2008

Valuation of Travel Time Reliability in Passenger Transport

ISBN 978 90 5170 846 2

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **439** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

Valuation of Travel Time Reliability in Passenger Transport

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Economische Wetenschappen en Bedrijfskunde
op maandag 24 november 2008 om 10.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Yin-Yen Tseng

geboren te Kaohsiung, Taiwan

promotoren: prof.dr. P. Rietveld
prof.dr. E.T. Verhoef

Acknowledgements

I am so happy to finally write the last page of this thesis which contains the results of my research undertaken at the Department of Spatial Economics of the Vrije Universiteit Amsterdam (VU) and the Tinbergen Institute (TI). I would like to take this opportunity to thanks the people that made the completion of this thesis possible.

First of all, I would like to express my deepest gratitude to my supervisors, Piet Rietveld and Erik Verhoef. I thank you for your enthusiastic supervision that always motivated me and helped me enjoy doing research, for your understanding in allowing me follow my own path and own ideas, for your encouragement in some of the difficult periods of my research, and for your confidence in me. It was my good fortune to have both of you as my advisors. I am especially indebted to Erik for providing his valuable time generously. I cannot remember how many times he replied to my urgent questions by email very late at night, and how often I interrupted his work by dropping into his office about some trivial questions on numerous occasions. All this help is truly appreciated.

Throughout these years, I was very lucky to have many opportunities to work on joint projects such as Rosebud, MD-PIT, Transumo, and the project undertaken together with Significance and John Bates. The experience of working with these research teams has been unique. It clearly shaped my ideas and improved my research in a distinctive way.

I would like to express my gratitude to the people who helped me with the data collection. Thanks to the MD-PIT team, Barry in particular, for providing me with such a rich and challenging data set. I would also like to thank Govert-Jan (ProRail), Freek (NS), and many of their colleagues for their help and constructive comments on the questionnaire about railway reliability. The online questionnaire would not have been possible without the help from Jeroen Knoef (NetQuestionnaire). Thanks also go to Gerard (Significance), Marco (Significance), and John Bates for the joint work on the in-depth interview questionnaire, and to Christiaan, Eveline, Ron, Olaf, my colleagues at the Department, who helped me so much with the in-depth interviews that inevitably had to be in Dutch. Hadewijch deserves a special mention, for translating my questionnaires nicely into Dutch with great patience and enthusiasm. The questionnaires had to be improved and revised so many times, but Hadewijch always kindly did the translation again and again without any complaint.

The atmosphere at the TI and at the Department of VU was always excellent. I am grateful to the TI and its staff, Marian and Nora, for providing help and assistance during my MPhil year at the TI. I would like to thank Maarten Lindeboom in particular for accepting me on the course at the TI, and more importantly, for introducing me to Piet and Erik. Thanks go to Ana, Andre, Dinara, Haris, Marcos, Jose, Rute and Sandra for sharing with me the struggle MPhil and many great and unforgettable moments in Amsterdam. Ana and Rute are two sweet friends who always offered me their unique warmth. Thanks also to Carla, Jiajia, Ernesto, Maria, Martijn, Michel, Pan Lei, Robert, Sebi, Simonetta, Tsechun, Wendy, and Yang Nan, who made my life more pleasant and enjoyable at the TI. Equally, I would like to thank Alex, Eva, Gert-Jan, Ghebre, Jasper, Laura, Moshe, Olaf, Paul, Roberto, Roos, Vanessa, and many other colleagues at the Department, for their support and encouragement, which were the most essential things I needed for completing the thesis. Romy deserves a special word of thanks for her warmest friendship and company. Of course, I will never forget Elfie, the best secretary that we could have ever had, for helping me so much in every respect.

I would also like to thank the members of my assessment committee: Prof. Gerard de Jong, Prof. Peter Nijkamp, Prof. Harry Timmermans, Prof. Mark Wardman, and Prof. Henk van Zuylen, for the time they spent on reading the draft of the thesis and their useful comments and suggestions. Thanks also go to Mrs Ellman for the language correction of the thesis. Her corrections were very detailed and useful. In addition, I thank Paul and Martijn for their generous time spent on the translation job of the Dutch summary.

During these years, apart from the scientific work, I also very much enjoyed skiing in Europe. Thanks go to my friends in Taiwan who skied with me throughout these years: Argoon, Chihyu, Emo, Kuowei, Jue, Ader, Kumo, Joe, Dan, and many others. Thanks for those wonderful moments you have brought to me, and for keeping on reminding me that there are things in life more exciting than doing research.

Finally, I am forever indebted to my parents and Huang Jian for their unconditional love and never-failing support. Huang Jian, I cannot imagine life without you. My gratitude and love to you are infinite.

Thanks,
Yin-Yen, September 2008.

TABLE OF CONTENTS

TABLES

FIGURES

CHAPTER 1

1 INTRODUCTION

1.1 TRAVEL TIME RELIABILITY

Traditional analysis of transportation projects considers travel time as the primary cost to the ravelers. Travel time savings often comprise the major benefit derived from many transportation projects. But, more recently, other aspects of travel such as the reliability have been getting more attention, as can be seen by the increasing number of studies coming out in this area. Qualitative and attitudinal studies have consistently found that travel time reliability is an important dimension in travelers' decision making (see, e.g., Prashker, 1979; Chang and Stopher, 1981; Small et al., 1999; Noland and Polak, 2002.). Therefore, understanding how reliability is valued is important because it provides insight into how policy makers can provide better transport options that match users' expectations. To be able to discuss the reliability of travel time and its importance, we have to first define what we mean by 'reliability' and how it can be matched to the travelers' perception of reliability.

In essence, travel time reliability is the certainty or stability of travel time of any particular trip under repetition, and as such, it is related to the statistical concept of variability[1] (Bates et al., 2001). Variability in travel times introduces uncertainty for travelers in that they do not know exactly when they will arrive at their destinations. It is important to clarify here that this variability represents the *randomness* in experienced travel times, and does not include the effects that are non-random (such as recurrent congestion effects) and are predictable by the travelers. In fact, if the travel time and its resulting delay can be expected, travelers can usually plan in

[1] In this thesis, the terms of travel time reliability and travel time variability are used interchangeably. In fact, travel time variability is a more precise term with respect to the analyses in the thesis. Nevertheless, in the research area of economic valuation of travel time variability, the majority used the 'value of (travel time) reliability' rather than the 'value of (travel time) variability'. To adhere to the convention, the 'value of (travel time) reliability' are used in this thesis, and the travel time reliability and travel time variability are therefore viewed as synonymous in the present context.

advance to avoid the costs of being late (or early). It is the variation around expected travel times that troubles travelers the most.

Clearly then, the unreliability of travel time is unattractive and undesirable to the transportation users. Nevertheless, the mechanism of how travelers perceive travel time unreliability, and how they change their travel behavior in response to it, is complex and may not be straightforward to model or understand. Some people just simply budget some extra travel time, and choose either to depart earlier or to accept late arrivals to cope with unreliability; some may ignore it altogether; and others consider travel time unreliability as an annoying feature and have an aversion to it over and above the dislike of its direct consequences.

Regardless whether travel time reliability may mean different things to different people, much work has been done in defining the objective measures of travel time reliability (see, e.g., van Lint and Zuylen, 2005). Two distinct types of travel time reliability measures are most commonly used in the empirical research of reliability valuation[2].

The first type of measure includes the standard deviation of travel time (Jackson and Jucker, 1981; Senna, 1994), percentile difference of the travel time distribution (usually the inter-percentile difference of the 50^{th}-80^{th} or the 50^{th}-90^{th}) (Lam and Small, 2001; Small et al., 2005; the coefficient of variation (Small et al., 1995), and the width of uncertain travel time distribution (i.e. the difference between minimum and maximum possible travel times) (Hensher, 2001), etc, all of which consider the dispersion of travel time distribution, but with different statistical viewpoints on describing the variation in travel times. For instance, the measure of percentile difference focuses on the right tail (longer than average travel time) of the travel time distribution, and therefore it focuses on the consequences of unexpected delays for the travelers. On the other hand, the standard deviation of travel time measures how wide the spread of travel times is about the mean, and both the right- and left-hand sides of travel time distribution are considered in this case.

---

[2] Here, we focus on the literature with respect to the valuation of travel time reliability for economic appraisal. In the research area of traffic management and transport planning, more measures of travel time reliability have been defined, such as the buffer time index, tardy trip measures, and probabilistic measures. See van Lint et al. (2008) for an overview.

The second type of measure focuses on schedule delay, defined to as the difference between a traveler's preferred and actual arrival time. Despite the use of 'delay', it can refer to a difference in either the early or the late direction, where the early side is called 'schedule delay early' and the late side is called 'schedule delay late'. The most common example is that of having a fixed work start time: if someone prefers to start work at 09:00 but actually arrives at 08:45, he incurs a schedule delay early of 15 minutes.

In applied work, these two types of measure are considered under different circumstances and by different modeling approaches. A measure such as the standard deviation is easier for practitioners who wish to quantify the amount of variation in travel times, but is not easily understood by non-statisticians, and, moreover, may not capture all elements of a distribution that matters to the travelers. On the other hand, the measure of schedule delay is probably often the best way to describe the effect of an individual's travel experience on decision making, since it conveys information about the major impacts of travel time unreliability. However, the computation of schedule delay requires information about the travelers' preferred arrival times, and thus cannot be directly obtained from the travel time distribution data. Although different in approach, the two methods may, of course, capture the same disutility. For example, for some travelers whose expected arrival time coincides with the desired arrival time, an increase in travel time variability would imply an increase in expected schedule delays. Nevertheless, these two types of measure are not necessarily mutually exclusive in modeling travelers' behavior response to the changes of travel time unreliability: unreliability may bring along disutility in addition to expected schedule delay costs.

1.2 WHY VALUE TRAVEL TIME RELIABILITY: MOTIVATION

The value of travel time reliability

Economics, as a science, deals with decisions under conditions of scarcity. As Lionel Robbins stated in his 1932 essay: "Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses." Scarcity is said to exist as long as there are opportunity costs involved in increasing the availability of a certain good. Though no market price can be directly observed, reliability is a scarce good: to make travel times more reliable, opportunity costs need to be incurred. And, to reduce the consequences of unreliability travelers may decide to depart earlier than desired, or accept late arrival. Such reserved additional travel time or late arrival is usually considered to be less productive and non-beneficial for travelers, and therefore there is an opportunity cost for reducing the consequences of travel time unreliability.

A number of empirical studies have demonstrated the importance of considering travel time reliability in the derivation of traveler cost functions (see, for example, Abdel-Aty et al., 1995; Small et al., 1995, 1999; Bates et al., 2001; Liu et al., 2004; Bhat and Sardesai, 2006.). These studies indicate that, depending on the circumstances, reducing the unreliability associated with trip times can offer significant traveler benefits. In essence, the value of travel time reliability is the amount that a traveler would be willing to pay in order to save one unit of travel time unreliability, and this willingness-to-pay can be obtained from travel demand models as the implicit tradeoff between reliability and money.

The motivation for valuing travel time reliability

A general objective in transportation is to achieve greater efficiency of mobility, and this is often done by, for example, shortening the average travel time, or enhancing travel time reliability, etc. A first motivation for valuing travel time reliability concerns the evaluation of major transport investments. These are usually subject to a cost-benefit analysis (CBA) within a standard framework, so that the policy makers can weigh and evaluate possible costs and benefits that are expected to be generated. The objective is usually to make (public) decisions in a more rational and transparent way, which should contribute to social welfare. Since the reliability of travel time is becoming increasingly important in policy evaluation, many Dutch infrastructure projects in the near future will focus not only on gains in average travel times but also on reductions in travel time unreliability (Hamer et al., 2005). Nevertheless, the current practice in estimating the time and reliability (delay)-related benefits in CBAs is usually only based on the evaluation of the reduction of average travel times. Failing to account for the benefits from reliability improvement could sometimes seriously limit the scope of the CBA, and may raise doubts regarding the accuracy of the results. For this reason, it is important that the monetary value of travel time reliability can be established, so that the benefits resulting from changes in travel time reliability can be incorporated into the standard CBA framework.

Apart from its use in investment evaluations, as a second motivation, the value of unreliability could also play a role in the determination of optimal road prices if unreliability varies with the level of road use, as is plausible. A similar argument may hold for public transport. Unreliability then constitutes an externality, like congestion, motivating marginal external changes that can be determined only when the marginal value of unreliability is known. And finally, the value of unreliability is an important input for the prediction of behavioral responses to transport policies (or other measures) that would affect reliability (see, e.g., Ettema et al. 2005, for the application in the traffic assignment model). So, ‘prediction’ is a third motivation for valuing reliability.

1.3 OBJECTIVES

The theme of valuing travel time reliability has become increasingly popular in the academic literature during the last decade. Many transport researchers have admitted the importance and necessity to incorporate the value of reliability into the CBA framework. Owing to the nature of the complex concept of reliability, there are still a fair number of unsettled issues in the research of this area. Our study aims to address some of these issues. The main objective of this thesis is to evaluate the monetary value of travel time reliability in the context of passenger transport. To evaluate the travel time reliability in a broadly accepted way, the value of schedule delay – a closely reliability-related measurement – will also be assessed. In addition to the empirical valuation of reliability, this thesis also aims to investigate a number of research questions, which aim to present an appropriate approach to fully address the theme of reliability valuation. The following research questions will be addressed:

1. Which factors influence estimates of values of reliability and schedule delay? (Ch. 2)
2. What is the best way of presenting reliability information in stated choice experiments? (Ch. 3)
3. a) What are the appropriate reliability and schedule delay estimates for various groups of travelers for different trip and individual characteristics, in the context of road and public transport? And b) Do different modeling approaches result in different estimates

in reliability and/or schedule delay variables? (both questions tackled in Ch. 4 and Ch. 6)

4. What is the impact of different utility specifications on the values of time and reliability? (Ch. 5)
5. a) How can schedule delay value estimates be applied to predict transit riders' behavior in anticipating departures, and to derive the value of unreliability as implied by schedule delay cost? And b) What is the impact of the naïve cost-benefit-analysis that fails to account for the benefit of reliability improvement?(Ch. 7)

1.4 METHODOLOGICAL APPROACH

In order to achieve the objectives and to answer the research questions mentioned above, several methodological approaches are followed. Here, we briefly discuss the approaches that will be employed in this thesis:

a) ***Meta-analysis:*** Meta-analysis is a quantitative form of research synthesis that aims to extract useful generalizations from a large body of diverse literature. It is done by combining the results of existing empirical studies, and then explaining the reasons for the differences in their results. This methodology has been well-established in the experimental sciences (see Cooper and Hedges, 1994), but is increasingly popular in economic analysis. As a starting point, we carry out a comprehensive literature review of the empirical research on the valuation of travel time reliability, and then exploit the meta-analysis technique in order to identify the systematic differences between the estimates of travel time reliability in empirical studies.

b) ***Valuation method:*** The standard practice for measuring the monetary values of travel time reliability is by estimating the coefficients of utility function containing the cost and reliability attributes in a discrete choice model, and then converting to monetary units by dividing by the cost coefficient. The data used for estimating a discrete choice model can be either revealed preference (RP) or stated preference (SP) data. The RP type of data relates travelers' choices observed from the real world, while the SP type of data comes from choices based on hypothetical situations. An extensive literature has contributed to the discussions of the relative advantages of these two types of data (see, e.g., Louviere et al., 2000). In determining the value of travel time reliability, however, the SP approach is favored by most researchers because it is difficult to find RP situations where there is sufficient variation in the reliability and other attributes (Bates et al., 2001). Since obtaining suitable RP data was not possible at the time when this research was carried out, collecting SP data became the only practical way in the context of this study. The empirical data used in this thesis will therefore be based on SP data, and more specifically, stated choice (SC) data.

c) ***Modeling travel behavior:*** The existing literature on modeling travel behavior for evaluating travel time reliability is well summarized in Noland and Polak (2002). According to their review, there are two competing modeling approaches in the literature: the mean-variance model and the scheduling model. These two models differ in the assumptions of how reliability is perceived and interpreted by the travelers. The scheduling model assumes that utility is affected by the traveler's scheduling consideration, i.e. the amount of time spent in being early or late, which is the consequence of travel time unreliability. The mean-variance model suggests that reliability is an additional term (usually the standard deviation or the variance of travel time distribution) besides the mean travel time in the utility function. In this thesis, we will mainly focus on these two approaches to model the travel behavior. Details of both models and the discussions of comparisons will be given the following chapters.

d) ***Theory of choice under uncertainty:*** The basic neoclassical economic theory of individual choice under uncertainty is Von Neumann and Morgenstern's (1944) expected utility theory. The decision rule to resolve the choice in an uncertain situation is called 'Maximum expected utility' (MEU), assuming that travelers would choose the alternative that maximizes the expected value of an appropriately defined utility. In the contemporary literature of behavioral economics, there is accumulating evidence showing that individuals' choices may violate the axioms of expected utility maximization (Kahneman and Tversky, 2000). The alternative theory to deal with choices under uncertainty is called prospect theory (Kahneman and Tversky, 1979), or its later extensions such as reference dependent theory and cumulative prospect theory. Recently, the application of prospect theory is gaining popularity in transport research to model traveler' choice behavior under uncertainty (see, e.g., Avineri and Prashker, 2005; Senbil and Kitamura, 2004). While it is important to acknowledge the progress in behavioral economic models, the MEU remains the standard tool to model travelers' choices (Batley 2007), and more importantly, the standard instrument to assess the value of reliability that can be applied in cost-benefit analysis. Therefore, we will adhere to Von Neumann and Morgenstern's MEU for behavior modeling in the context of this thesis.

1.5 OUTLINE OF THE THESIS

The structure of this thesis is shown schematically in Figure 1.1. The analysis starts with an overview of the empirical studies on reliability estimates. It appears that there is a great deal of variety in these estimates. On the basis of the review, we identify the sources of the variations and perform a meta-analysis to explain the differences of these estimates in a systematic way. This is done in Chapter 2. In Chapter 3, we review how reliability information can be presented in previous stated preference studies. Based on different presentation formats, an in-depth interview survey is conducted. The purpose of the in-depth interview is to see how respondents perceive and interpret travel time reliability, and to test their understanding of different ways of presenting reliability. Based on the results of these interviews, various formats of reliability presentation will be recommended in this chapter.

The next two chapters present the empirical findings of an extensive stated choice experiment survey conducted among Dutch car commuters who regularly experience congestion. In this survey, the respondents faced a choice experiment in which they distributed 10 trips over 4 constructed alternatives. The estimations report the results for values of time, schedule delay, and uncertainty in the context of car transport; both from conventional multinomial logit models and from mixed logit models. The latter are considered to resolve the problem of repeated choices existing in the data set. In Chapter 4, the analysis not only focuses on the central estimates but also investigates the factors, e.g. trip and individual characteristics, that influence these estimates. In Chapter 5, an alternative modeling framework is proposed to incorporate commuters' strong time-preferences over the morning peak. This alternative model allows us to estimate time-dependent values of travel time savings and schedule delay, which follow plausible and intuitive patterns.

Chapter 6 presents the empirical estimates of a stated choice experiment survey of Dutch railway passengers. Similar to the analysis in Chapter 4, the central estimates and the estimates for various groups of respondents are reported. In the context of public transport, it is common for people to choose to take an earlier than necessary connection when the service is unreliable, in order to avoid the consequences of possible delays. Chapter 7 investigates the behavior of such “anticipating departures” for a scheduled public transport service, under the presence of unreliability. Furthermore, the value of implied schedule delay, i.e. the marginal expected schedule delay costs associated with a change in the standard deviation of travel time distribution, will also be derived for some specific cases.

Finally, Chapter 8 gives the conclusions of this thesis.

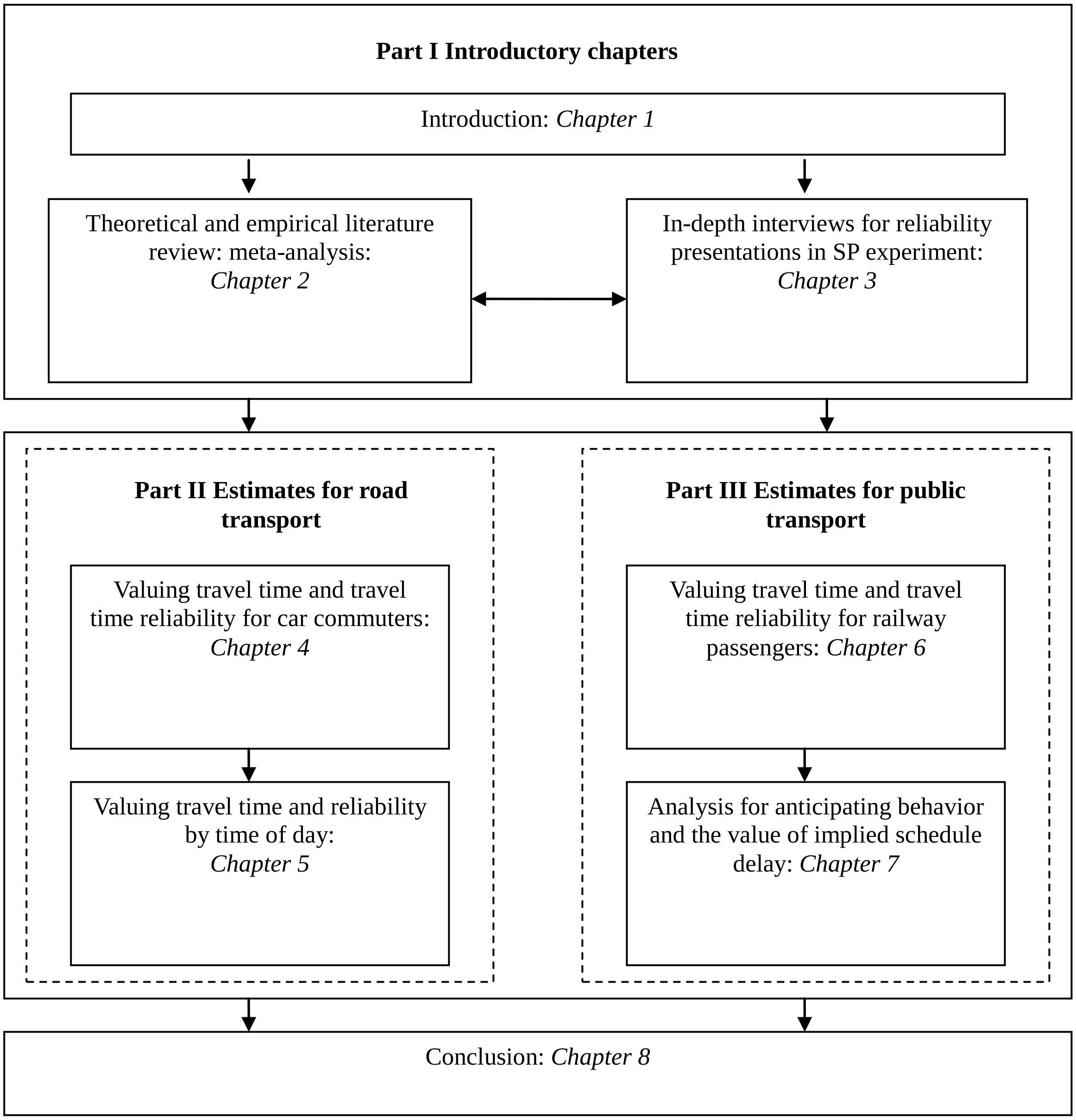


Figure 1.1 Thesis layout

CHAPTER 2

2 A META-ANALYSIS OF VALUATION OF TRAVEL TIME RELIABILITY

2.1 INTRODUCTION

Various factors govern travel behavior. Along with travel time, reliability of travel time is increasingly being regarded as an important component in an individual's decision making concerning route choice or mode choice (e.g. Small and Verhoef, 2007). The concept of 'reliability' reflects that an individual has to make his or her travel decision under uncertain circumstances with respect to travel time, and hence the nature of reliability can be described by the distribution of travel time (Bates et al., 2001).

During the last decade, a fair number of studies have attempted to incorporate travel time reliability attributes into travelers' choice models. However, there is still debate on reliability valuation, particularly with respect to the way of measuring travel time reliability and model specification; and point estimates differ substantially between studies. No consensus seems to have been achieved thus far, either on point estimates or on the methodological question of how to measure the value of reliability.

In this study we review empirical estimates of values of reliability of travel-time-related attributes. We not only look into the valuation of travel time reliability itself but also consider the valuation of scheduling delay variables. Our aim is to study the sources of variation in empirical estimates, and to investigate some unresolved issues by means of meta-analysis, a quantitative method for literature survey. By performing a meta-regression, differences in estimates can be identified and explained in a systematic way. Such a meta-regression analysis can also offer a prediction of the results that such new research may find, and thus suggests potentially fruitful lines for future inquiry (Stanley, 2001).

The chapter is organized as follows. Section 2.2 considers the concepts of value of time, reliability, and scheduling cost. It also shows the most commonly used empirical modeling approach to travel time reliability valuation. Section 2.3 discusses the main sources of disagreement and possible sources of variation in empirical works. Section 2.4 describes the data and gives an overview of empirical estimates of various reliability indicators, including the reliability ratio, scheduling delay early ratio, and scheduling delay late ratio. The meta-regression results and discussions are presented in Section 2.5, and Section 2.6 concludes.

2.2 THEORETICAL FRAMEWORK

2.2.1 Empirical model

The conventional approach to modeling travelers' choice behavior is discrete choice analysis, which stems from utility maximization theory and assumes that respondents will select the alternative in a given choice set that has the highest utility. Among the various models used in discrete choice analysis, the *random utility model* (RUM) is the most intensively used in the empirical assessment of travel behavior. In such an approach, the utility of individual i from choosing alternative j is given by (Ben-Akiva and Lerman, 1997):

$$U_{ij} = V_{ij} + \varepsilon_{ij} . \qquad (2.1)$$

The first part V_{ij} of Eq. (2.1) is the 'deterministic part' or 'systematic part', and is determined by the observed attributes of the alternative and the characteristics of the individual. The most basic model is the linear additive form, represented as:

$$V_{ij} = \sum_k \beta_{ik} x_{ijk} , \qquad (2.2)$$

where subscript k represents the set of attributes that may affect the individual's utility in choosing alternatives j.

The second part ε_{ij} of Eq.(2.1) is the random (or error) term, which is unobserved by the researcher. Various models can be derived from different assumptions on the error term distribution. In practice, the most popular are models from the logit family, which assume that the error term follows an extreme value type 1 distribution. The advantage of the logit model is its tractability, though it imposes restrictions on the covariance structure of error terms. Therefore, many models have been developed that deviate from the standard logit, such as the nested logit and the mixed logit, which aim to relax the restrictions on error terms.

2.2.2 Concepts of reliability and scheduling delay

Since the concept of travel time reliability is closely connected with the distribution of possible travel times, at least two dimensions of possible travel times have to be considered in modeling the effect of reliability: namely, the possible magnitudes and probabilities with which these occur. Besides the expected travel time measurement, the indicator that can reflect the distribution of travel time is usually the variance or the standard deviation of travel times.

Along with the utility loss incurred by the unreliability in travel time, a traveler may also attach some disutility to arriving at the destination before or after some preferred arrival time (PAT). Thus, the (expected) difference between actual arrival time and preferred arrival time may play a role in traveler's decision making. Following Small (1982), this difference between the PAT and the actual arrival time can be called schedule delay (SD): $SD = PAT - [t_h + T(t_h)]$, where t_h is the departure time, and the amount of travel time $T(t_h)$ depends on the chosen departure time. People may value early and late arrivals differently according to the different consequences. Most studies (see, e.g., Small, 1982; Bates et al., 2001) therefore evaluate SD as two separate terms, schedule delay early (SDE) and schedule delay late (SDL), which can be expressed as: $SDE = \text{Max}\,(0, PAT - [t_h + T(t_h)])$ and $SDL = \text{Max}\,(0, [t_h + T(t_h)] - PAT)$.

2.2.3 Modeling approaches

The earliest work to consider the effect of reliability in travel behavior used the mean-variance approach. Jackson and Jucker (1981) specified a model where a traveler can make the tradeoff between travel time and variance of travel time explicitly. Both of these two elements were included in a cost function that travelers seek to minimize. A general form of this mean-variance approach is given by Eq. (2.3):

$$Min\ C = E(T) + \lambda \cdot Var(T) \tag{2.3}$$

The coefficient of variance of travel time λ can be seen as a true monetary value for unreliability (if travel cost is included in the functional form, then λ requires division by the cost parameter to become a true "value": see below). Instead of $Var(T)$ in Eq. (2.3), sometimes the standard deviation has been used. This mean-variance approach, originally developed in the field of portfolio analysis in financial markets, can easily be applied in mode or route choice analysis. Yet the weakness of this approach might be its inability to deal with departure time choice behavior with scheduling constraints, to which we will now turn.

As mentioned, Small (1982) introduced the schedule delay (SD) variable to measure the difference between a traveler's actual arrival time and preferred arrival time (PAT). Since people may value early and late arrivals differently because of the different consequences, the SD variable can be evaluated as two separate terms, schedule delay early (SDE) and schedule delay late (SDL). SDE is defined as the amount of time arriving earlier at the destination than the PAT, while SDL is the amount of time arriving later than the PAT. A linear indirect utility function that captures both could be:

$$U = \beta_T \cdot T + \beta_C \cdot C + \beta_\beta \cdot SDE + \beta_\gamma \cdot SDL + \beta_\theta \cdot D_L\ , \tag{2.4}$$

where T denotes the travel time, and C gives the monetary travel cost. SDE is defined as Max(0, PAT minus actual arrival time), SDL is defined as Max(0, actual arrival time minus PAT), and D_L is the lateness dummy, which is equal to 1 when $SDL \geq 0$, and 0 otherwise. The coefficients of β_β and β_γ measure the disutilities of being early and late, while β_θ represents a fixed penalty for late arrival. Since T, SDE and SDL are disutilities, the coefficients are assumed to be negative. Small's (1982) empirical finding is that $|\beta_\gamma| > |\beta_T| > |\beta_\beta|$, which means that people prefer early arrival to additional travel time, and prefer additional travel time to late arrival.

The model proposed by Noland and Small in 1995 extended Small's 1982 trip-scheduling model of Eq.(2.4) by considering the probability distribution of travel time. The result is presented as Eq.(2.5), which is based on *Maximum Expected Utility* (MEU) theory:

$$\mathrm{E}(U) = \beta_T \cdot \mathrm{E}(T) + \beta_C \cdot C + \beta_\beta \cdot \mathrm{E}(SDE) + \beta_\gamma \cdot \mathrm{E}(SDL) + \beta_\theta \cdot P_L , \tag{2.5}$$

where $\mathrm{E}(T)$ is the expected travel time; $\mathrm{E}(SDE)$ is the expected schedule delay early; $\mathrm{E}(SDL)$ is the expected schedule delay late; and $P_L \equiv \mathrm{E}(D_L)$ is the lateness probability.

The model shown in Eq.(2.5) assumes that scheduling considerations only result from disliking arriving too early or too late, as there is no other reliability-related variable added in this model. Although this hypothesis could be true in some cases, it is more likely that people have some additional inherent aversion to unreliability, due to factors such as inconvenience, stress, or anxiety. To account for all these ill effects of unreliability, some studies (e.g. Small et al., 1999) add the variable of standard deviation of travel time (*STD*) to Eq.(2.5), so that the additional disutility of unreliability can also be captured in the model, apart from the scheduling considerations. The model specification is shown in Eq.(2.6):

$$\mathrm{E}(U) = \beta_T \cdot \mathrm{E}(T) + \beta_C \cdot C + \beta_\beta \cdot \mathrm{E}(SDE) + \beta_\gamma \cdot \mathrm{E}(SDL) + \beta_\theta \cdot P_L + \beta_R \cdot STD . \tag{2.6}$$

Note that the specification in Eq.(2.6) implies the combination of the scheduling model and the mean-variance approach. Most of the empirical studies on reliability valuation have estimated the models based on Eq.(2.3), Eq.(2.5), or Eq.(2.6). Our main interest of analysis in this study will therefore be the parameters of reliability[3], schedule delay early, and schedule delay late, all compared with the parameters of the travel time or cost term.

Once the model is estimated, one can find the marginal rates of substitution between any pair of the attributes in the bundle. The monetary value of travel time (VOT) is defined as the marginal substitution rate between travel time and costs, and hence as the ratio of the respective coefficients:

$$VOT = \frac{\partial U / \partial T}{\partial U / \partial C} = \frac{\beta_T}{\beta_C} . \tag{2.7}$$

[3] Several measurements, such as standard deviation, variance, percentile difference of travel time, have been used by researchers to represent the variable of travel time reliability. We will discuss the reliability measurement issue in further detail in Section 4.2.

The monetary values of reliability (VOR), schedule delay early (VSDE) and schedule delay late (VSDL) can be expressed analogously.

One practical issue in the meta-analysis that will follow is that some studies do not include the cost-related terms in their estimated model. To increase the number of observations in the database, we therefore decided to use the marginal rate of substitution between time and reliability for our variable of interest in the meta-analysis. The marginal rate of substitution between travel time and reliability is called the *reliability ratio*, i.e. $RR = \beta_R / \beta_T$, defined by Black and Towriss (1993). To facilitate the empirical analysis of scheduling variables, we also define the *schedule delay early ratio* and the *schedule delay late ratio* as $SDER = \beta_\beta / \beta_T$ and $SDLR = \beta_\gamma / \beta_T$, respectively. Another advantage of using the reliability or scheduling ratio in the analysis is to avoid transformation problems because of exchange rates and price index developments.

2.3 Issues in the valuation of reliability

2.3.1 Revealed versus stated preference

Discrete choice models can be estimated on the basis of revealed and stated preference data. Traditionally, empirical studies of traveler's choice behavior rely on data from observing what people *actually* do, i.e. revealed preference (RP) data. However, recent studies use data from people's self-reported choice under hypothetical situations, which we refer to as stated preference (SP) data. As Louviere et al. (2000) summarized, there are some compelling reasons for economists to use SP data: for example, to estimate demand for new products with new attributes or features; to have sufficient variation of the explanatory variables to allow for reliable model estimation; or when observed explanatory variables are highly collinear in the marketplace. This last reason is the most common limitation in RP data, and may cause severe identification problems in econometric analysis.

The most serious critiques of SP data are probably its lack of reality, and doubts concerning the validity of hypothetical choices. However, many researchers believe that this problem can be solved, or at least reduced to acceptable proportions, by well-designed SP surveys. Indeed, well-designed SP data may be superior to poor RP data, which are problematic in model estimation. In the research area of reliability valuation, Bates et al. (2001) argued that SP is almost the only feasible method of data collection, because it is hard to obtain suitable RP data with sensible trade-offs between the travel time, reliability, and cost. Nevertheless, there are a few VOR studies that do have good quality RP data (Lam and Small, 2001; Ghosh, 2001; Yan, 2002) which were obtained from the loop-detectors in the time-varying tolling experiments. In any case, our interest in the present study is not to argue which method is most appropriate to use, but instead to see whether there is a systematic difference in the estimates between SP and RP data (see also Brownstone and Small, 2003).

Earlier studies (Ghosh, 2001; Yan, 2002) show that the median SP estimates of VOT and VOR are about half of the median RP estimates, and the differences are statistically significant. Similarly, a recent meta-analysis of value of travel time savings (VTTS) (Shires and de Jong, 2008) also finds that, for commuting and other passenger travel, SP and the joint SP-RP studies produce significantly lower VTTS than studies based purely on observed outcomes. Brownstone and Small (2003) hypothesize that the difference between SP and RP may be caused by the misperception of 'true' travel time losses in RP conditions: people may exaggerate the amount of delay time because of impatience with heavy traffic.

A potential issue for the VOR SP studies concerns how to present the reliability attribute in the SP experiments. As will be reviewed in Chapter 3, there is little consensus about the preferred presentation format that should be used in the SP experiment. It is likely that respondents may perceive the information about reliability (i.e. the distribution of travel time) differently for different reliability presentation formats. However, though it would be interesting to see if there is any pattern between the reliability estimates and the means of presentation, the existing data does not allow us to do this task. This is because almost no study adopted exactly the same presentation, and, consequently, it is difficult to group the similar studies together and to make comparisons between groups. Therefore, in the present paper, we will only focus on investigating the differences of reliability estimates between RP and SP data. The issue of reliability presentation format in the SP experiment will be further discussed in detail in Chapter 3.

2.3.2 Utility specification: reliability versus scheduling variables

Two UK studies: namely Arup Transport Planning (2002) and Bates et al. (2003), concluded that the value of reliability can be entirely explained by the implications for the expected scheduling cost. Indeed, some empirical studies (e.g. Small et al., 1995; Small et al., 1999) obtained insignificant coefficients for the effect of reliability when including both reliability and expected scheduling variables in the model. This might reflect that, besides the effect upon the expected scheduling costs, unreliability causes no significant inconvenience in itself. Another possibility could, of course, be measurement problems with schedule delays, unreliability, or both.

Though the concepts of reliability and scheduling are closely related to each other, they should not be treated as identical. The idea is that, apart from people's scheduling preference, they may have some additional disutility on account of the inconvenience or anxiety caused by unreliability of travel time, even when the 'expected scheduling delay' cost is the same. Moreover, a great number of trips do not have strict scheduling constraints (e.g. shopping or leisure), and people may be indifferent as long as they arrive at the destination within a certain range of arrival times. In such a case, the disutility may come from the inconvenience of planning due to the unreliability of travel time rather than scheduling considerations. One of our main purposes is to investigate this utility specification effect, and to see what the extent of this influence is.

2.3.3 Types of choice set

Another question is whether the estimations of reliability and scheduling variables differ over types of choices (e.g. route choice, mode choice, departure time choice.). Route choice problems are also referred to as 'within-mode' choices, as opposed to 'between-mode' choices, whereas the departure time choice can be incorporated into both types of choices.

Basically, if the underlying utility function is correctly specified to reveal traveler's actual choice behavior, the estimates of reliability and scheduling variables obtained from different choice set domains should be close to each other. However, some concerns may arise in practice. For example, since not all the alternatives are available to the respondents in the real mode choice problem, observed behavior might differ from the hypothetical behavior.

One of our aims in this study is to investigate whether the studies in our data set reveal substantial differences in valuation between the different types of choice sets, that is, between-mode choice and within-mode choice.

2.3.4 Trip purposes and modes

There is an enormous literature on empirical estimates of the value of time (VOT) based on different trip purposes and mode used. A general belief is that the VOT varies with these trip characteristics. Wardman's (2001) meta-analysis tracked the effects of various trip attributes on VOT and found that there are considerable differences across modes and trip purposes. Consequently, one may expect that these trip characteristics have similar effects on the value of reliability. This issue will be examined in our meta-analysis by looking at whether there are systematic variations between commuting and other trips, as well as between public and private transport.

2.3.5 Individual characteristics: observed and unobserved heterogeneity

There is considerable evidence that the VOT varies with individual characteristics such as income and gender (see, e.g., Small et al., 1999; Lam and Small, 2001; Wardman, 2001.). To investigate the variations on the estimates, one can specify the 'observed' covariates in the model, or treat the source of variations as 'unobserved' (by randomizing the parameters and/or allowing more general correlated error structure forms), or have both of 'observed' and 'unobserved' heterogeneity in the model.

Observed heterogeneity in the estimates can in practice be evaluated by incorporating the interaction terms of those trip or individual traits variables with travel time, reliability, or cost variables. Whilst the idea of testing whether these covariates have important effects in valuation seems interesting, our data do not allow us to do this test. The reason is that almost no study included the same set of covariates. For example, Small et al. (1999) considered the effects of income and household composition on the valuations of reliability and schedule delay variables, while Lam and Small (2001) investigated the gender effect on the value of reliability. Because each study had its own interest and purpose in exploring this issue, it is therefore difficult to compare the estimates of specific groups with each other directly.

More recent studies have taken unobserved heterogeneity into account, thanks to the advances in econometrics modeling techniques and computing power. In the literature (Hensher, 2001; Hensher and Greene, 2003), there are two considerations to accommodate the unobserved variability of preferences into the model: (a) allowing correlation structures of error terms; and (b) randomizing the parameters associated with each attribute. It is not clear whether incorporating unobserved heterogeneity will lead to upward or downward adjustments of values. Hensher (2001) suggested that the less restrictive choice model tends to produce higher estimates, while Ghosh (2001) showed that the most general model yielded the lowest estimates, which contradicts Hensher's results.

We aim to investigate the effect of accounting for unobserved heterogeneity on the reliability and scheduling ratios, our dependent variables. Because different degrees of complexity were specified in each study in order to take account of the unobserved variability, it is hard to categorize studies, either according to the exact way of randomizing parameters or allowing for sophisticated error structures. Thus, we only consider the effect on estimates of accommodating unobserved heterogeneity in any such sense.

2.3.6 Different measurement in attributes

There are various measurements of reliability in empirical assessments, such as the standard deviation, the coefficient of variation, the difference between the 90^{th} and the 50^{th} percentile of travel time, etc. This lack of consensus on how to characterize the reliability by a common variable creates a problem when comparing empirical estimates, and this issue will be discussed in more detail in Section 2.4.2.

In addition to the wide range of reliability measurements, travel time is also expressed in different ways, such as mean or median travel time, free-flow time, congested time, and median delay time. Since the value of time is the denominator of reliability ratio, these different measurements in travel time may have an influence on our variable of interest. In particular, previous studies have indicated that the value of congested time is considerably higher than the value of free-flow time or uncongested time (Hendrickson and Plank, 1984), and delay time is evaluated higher than in-vehicle travel time (Wardman, 2001). Thus these different measurements of travel time may also play a role in the variation of reliability ratios and schedule delay ratios.

However, if we want to classify each travel time measurement in different categories, we would have very few observations in certain categories. Therefore, in order to solve this problem, we select some of the conceptually similar measurements, e.g. 'congested travel time', 'median travel time savings' and 'mean delay', and then place them into the same group called 'congested travel time'. In the following analyses, this group will be compared with the group 'uncongested travel time', which is a combination of the 'travel time' and 'free-flow travel time' measurements.

2.4 METHODOLOGY

2.4.1 Data and sampling

To search the empirical estimates for reliability and scheduling variables, we started from the EconLit database, transportation research journals, and the Google search engine, including published papers, reports, and working papers[4]. A recent summarized study of travel time reliability from RAND Europe (2004) serves as a good reference for collecting the reliability estimates. We computed the reliability and scheduling ratios as explained at the end of Section 2.2.3. However, we excluded some estimates, which used diverging definitions of reliability and cannot be made comparable to other estimates (e.g. König and Axhausen, 2002[5] and Cascetta and Papola, 2003[6]). The study by Hendrickson and Plank (1984) is excluded because of the quadratic terms of scheduling variables specification in the choice model and the resulting exceptionally large value of the VSDL ratio. We also decided to drop Wilson's (1989) study in our meta-analysis, because of its extreme value of the VSDE ratio and the unclear definition of schedule delay variables. The overall studies and computed ratios are summarized in Table 2.1.

2.4.2 Making reliability estimates comparable

As mentioned in Section 2.3.6, there are various measurements of reliability, and these different uses of reliability measurement certainly create a problem of comparison (see Table 2.2).

---

[4] Since our variables of interest are the reliability ratio and the scheduling ratios, we only considered empirical studies that include the valuation of either both travel time and reliability or both travel time and scheduling variables.

[5] König and Axhausen used two separate variables 'duration of delay' and 'probability of delay' to present the effect of reliability.

[6] The early and late penalty variables defined by Cascetta and Papola are not equivalent to the schedule delay early and late variables defined in Section 2.2.2.

Table 2.1 Overview of studies with empirical estimates

Authors	Study type	Year of Publication	VOR ratio (RR)		VSDE ratio (SDER)		VSDL ratio (SDLR)	
			obs	mean	obs	mean	obs	Mean
Small	RP	1982	-	-	2	0.667	2	2.139
Lam and Small	RP	2001	25	1.137	6	0.456	4	0.762
Small et al.	SP	1999	3	2.515	4	0.867	4	5.010
Ghosh (Dissertation)	SP&RP	2001	5	0.986	-	-	-	-
Yan (Dissertation)	SP&RP	2002	30	1.082	-	-	-	-
Small et al.	SP	1995	3	0.536	4	0.872	4	1.813
Koskenoja (Dissertation)	SP	1996	7	0.408	7	0.507	4	1.396
Bates et al.	SP	2001	-	-	1	0.442	1	0.897
Hensher	SP	2001	6	0.574	-	-	-	-
de Palma et al.	SP	2003	-	-	12	0.487	12	1.494
de Jong et al.	SP	2003	-	-	9	1.059	9	1.422
Rietveld et al.	SP	2001	1	1.400	-	-	-	-
Liu et al.	RP	2004	1	1.610				
Hess et al.	SP	2007	-	-	24	0.870	24	1.440
Bhat and Sardesai	SP&RP	2006	3	0.492				
Hollander	SP	2006	1	0.100	-		1	2.769

Table 2.2 Different definitions of reliability used in empirical estimations

Reliability definitions	Notation	# obs	Min	Max	Mean
Standard deviation of travel time	STD	7	0.1	3.222	1.552
Coefficient of variation of travel time	CV	9	3.279	14.392	7.935
Difference between 90th and 50th percentile travel time	90DMP	29	0.483	1.714	1.131
Difference between 80th and 50th percentile travel time	80DMP	20	0.968	1.952	1.476
Min-Max	MM	9	0.268	1.152	0.546
Incident	INC	11	0.380	0.441	0.415

If we estimate the utility function for a given set of observations, the estimated coefficient for the standard deviation (STD) of travel times will not be equal to the estimated coefficient of the coefficient of variation (CV), i.e. $\beta_1 \neq \beta_2$ in Eq.(2.8). The ideal way to correct these coefficients based on different measurements is to go back to the original survey data, and then re-estimate the model again by using the standard definition of reliability. However, this is not feasible in our case. The second-best way to adjust these coefficients is by looking at the relationship between those different measurements, and then correct the coefficients according to these transformed relationships:

$$\begin{cases} U = \alpha \cdot \mathrm{E}(T) + \beta_1 \cdot STD + ... \\ U = \alpha \cdot \mathrm{E}(T) + \beta_2 \cdot CV + ... \end{cases} \tag{2.8}$$

Take the STD and the CV for example, we know in advance that there exists a relationship between the STD and the CV, that is, *CV*=*STD*/(mean travel time). Thus, we can infer that $\beta_2 = \beta_1 \times (\text{mean travel time})$.

Next, we can investigate the relations between the standard deviation (STD), the difference between the 90th percentile and median travel time (90DMP), and the difference between the 80th percentile and median travel time (80DMP) under various types of distributions. In the case of a uniform distribution, we can derive the analytical solutions for the relations between the STD, the 90DMP and the 80DMP. This shows that the values of the 90DMP and the 80DMP are proportional to the standard deviation. Thus, assuming that travel time follows a uniform distribution, we can easily correct the estimated coefficient of the 90DMP to the STD, based on the calculated ratio. A similar situation also holds for triangular distributions. For the normal distribution, since the analytical solution is difficult to implement, we use simulations to infer these ratios[7]. All the "transformation ratios" between these variables are listed in Table 2.3 for these three distributions.

Table 2.3 Transformation ratios between the STD and the 80/90DMP under various distributions

	Uniform distr.	**Normal distr.**	**Triangular distr.**
STD	1.000	1.000	1.000
90DMP	1.384	1.282	0.993
80DMP	1.038	0.842	0.661
Min-Max	3.464	-	4.899

From Table 2.3 we find out that the values of the transformation ratios of the normal distribution are located in-between the values of uniform and triangular distributions. Therefore, we decided to choose the transformation ratios for the normal distribution as our "correction factor". We hypothesize that the distribution of travel time is normally distributed, and then correct the reliability estimates to make them comparable.

In the case of the Min-Max, defined as the difference between minimum and maximum possible travel times (MM), although it is true that the concept of the MM is relatively different from STD and 80/90DMP (from a statistical point of view), we seek to convert these estimates under some reasonable assumption. The bottom row of Table 2.3 shows the transformation ratios between the MM and the STD for uniform and triangular distributions[8]. Here we choose uniform as the

---

[7] The ratio between the STD and the 80/90DMP is constant in the normal distribution. In other words, the ratio is independent of the standard deviation in the normal distribution.
[8] With the assumption that the MM equals the range of upper and lower bounds in the uniform and triangular

hypothetical distribution for the MM estimates correction. Nevertheless, we will add an explanatory dummy in the meta-regression to see if there is a substantial difference between this group of estimates and others.

The correction approach described above is used in correcting estimates between the STD, the CV, the 90DMP, the 80DMP, and the MM. Unfortunately, it is not possible to follow the same procedure for the estimates of ‘incident’ under any reasonable assumption. We therefore decided to drop these reliability estimates from our formal meta-analysis.

2.4.3 Overview of empirical estimates

Before performing the meta-analysis on our sampled estimates, we first look at the descriptive statistics and the conditional means of the dependent variables in the meta-regressions. The descriptive statistics of RR, SDER, and SDLR are presented in Table 2.4. The values of RR here have been adjusted according to the correction procedure described in Section 2.4.2. The sample means of SDER and SDLR imply that the general belief that $VSDE < VOT < VSDL$ basically holds in our sampled observations.

Table 2.4 Descriptive statistics of adjusted RR, SDER, and SDLR

	VOR ratio (RR)	VSDE ratio (SDER)	VSDL ratio (SDLR)
Mean	1.325	0.745	1.652
Median	1.349	0.748	1.174
Standard deviation	0.682	0.397	1.392
Minimum	0.100	0.166	0.164
Maximum	3.990	2.460	7.150
observations	74	69	67

Next, Table 2.5 gives the conditional means of RR, SDER, and SDLR for various sub-groups within the sample. The conditional means of RR on choice types, modes, and trip purposes are absent because of small subsamples (only 1 or 2 observations). Serving as the preliminary stage of meta-analysis, these conditional means give a rough idea of how these factors affect the variables in which we are interested. As we can see from Table 2.5, these conditional means vary

distributions, we can derive the transformation ratios for these two distributions.

significantly in several between-group comparisons, such as trip purpose, travel time and reliability measurement, and utility specification. Some findings in Table 2.5 confirm our expectations: for instance, including both reliability and scheduling variables produce lower estimates of both variables.

In the next section, we will explore the data more systematically in the meta-regression.

Table 2.5 Conditional means for various categories of studies

	VOR ratio (RR) (n=74)		VSDE ratio (SDER) (n=74)		VSDL ratio (SDLR) (n=70)	
Groups	n	Mean	n	Mean	n	Mean
Data Types						
Revealed preference	48	1.367	8	0.509*	8	1.104
Stated preference	26	1.248	61	0.776*	59	1.726
Choice types						
Between-mode choice	-	-	33	0.926***	34	1.475
Within-mode choice	-	-	36	0.579***	33	1.834
Mode specific estimate						
Private transport	-	-	49	0.787	48	1.830*
Public transport	-	-	20	0.641	19	1.200*
Trip purpose						
Commute	-	-	46	0.711	44	1.985***
Others	-	-	23	0.812	23	1.014***
Unobserved Heterogeneity						
Not accounted for	63	1.281	33	0.575***	30	1.957
Unobserved hetero.	11	1.579	36	0.901***	37	1.404
Travel time measurements						
Uncongested travel time	44	1.260	54	0.824***	52	1.813*
Congested travel time	30	1.421	15	0.461***	15	1.094*
Reliability measurements						
Min-Max	9	1.893***	-	-	-	-
Others (standard deviation, percentile difference)	65	1.247***	-	-	-	-
Utility specification I						
No scheduling/reliability variable	59	1.458***	54	0.805**	55	1.766
Including scheduling / reliability variable	15	0.805***	15	0.528**	12	1.130

Note: The statistical test (t-test) is concerned with the comparison of means within each group. Significance is indicated by ***, **, and *, referring to significance at the 1%, 5%, and 10% level, respectively.

2.5. META-REGRESSION VARIANTS AND RESULTS

In this chapter we employ a meta-regression to investigate whether there is systematic variation in the effect sizes that is caused by the study characteristics and other conditioning variables that we discussed in Section 2.3. The variables used in the meta-regressions are listed in Table 2.6. In performing the meta-regression, one fundamental aspect is the treatment of standard errors of the effect sizes. A common practice in meta-analysis is to weight each effect size by the inverse of its standard error, in order to give a higher weight to a more precise estimate. However, the estimated standard errors of the effect sizes are only available for around half of the observations in the primary studies[9]. Since the sample sizes of the underlying studies can serve as a proxy to account for the precision of the effect sizes (see, e.g., Florax et al., 2005), we used the information of the sample sizes in the primary studies to calculate the appropriate weights.

Table 2.6 Meta-regression variables

Dependent variables	RR	SDER, SDLR
Explanatory variables		
Data type	*SP* = 1 if SP data is used	*SP* = 1 if SP data is used
Choice type	-	*BET* = 1 if between-mode choice is used
Unobserved heterogeneity	*HET* = 1 if unobserved heterogeneity is accounted for	*HET* = 1 if unobserved heterogeneity is accounted for
Trip mode	-	*PUB* = 1 if public transport
Trip purpose	-	*COM* = 1 if commuting trip
Utility specification	*SD* = 1 if schedule delay variable(s) is added in the utility	*R* = 1 if reliability variable is added in the utility
Travel time measurement	*TT_con* = 1 if congested travel time is used as the travel time variable	*TT_con* = 1 if congested travel time is used as the travel time variable
Reliability measurement	*MM* = 1 if Min-Max is used as reliability variable	-

9 Because some studies specified quadratic terms for travel time, reliability or scheduling variables, or used covariates (e.g. distance) in the choice models, we therefore used the 'mean' VOT, VOR, VSDE, or VSDL provided from the authors in the summary results. In some studies with mixed logit models, we also used the (simulated) 'mean' VOT and VOR reported by the authors, rather than deriving the effect sizes directly from the reported coefficients. In these cases, it is not possible to obtain the standard errors of the effect sizes from the original estimated coefficients of the models.

Tables 2.7-2.9 summarize the results of the meta-regression of the RR, the SDER, and the SDLR, respectively. The included sets of explanatory variables aim to investigate those sources of variation discussed in Section 2.3. Yet not all the factors can be investigated in our present study, because of collinearity between the explanatory variables. Nevertheless, the large part of the main sources of variations in the RR, SDER and SDLR estimates can still be investigated in this framework. Column I in Tables 2.7-2.9 shows the meta-regression results estimated by means of weighted least squares (WLS) with weights equal to the sample size (i.e. the proxy for the standard errors of the effect sizes) of the original studies. In the meta-analysis literature, the WLS specification is often referred to the ‘fixed effects’ model (Hedges, 1994), assuming that the heterogeneity of the effect sizes is systematic and can be identified by the differences between studies. In this type of model, the source of the variation of the effect sizes comes from the sampling error only.

Another attractive specification is the ‘mixed effects’ model. In the mixed effects model, the heterogeneity of the effect sizes is partly systematic (derived from identifiable differences between studies) and partly random, and the effect sizes from different studies are assumed to be randomly drawn from a normal distribution. The mixed effects model is generally preferred if part of the heterogeneity can not be captured by the explanatory variables in the model. The results of the mixed effects model are shown in Column II of Tables 2.7-2.9.

The preceding two models assumed that every effect size from the primary studies is of equal importance, regardless of its quality and origin. However, some authors tended to publish only one of their favorite estimates in a study, while some tended to publish multiple plausible specifications[10] so that many estimates were produced in a study. To reduce the impact of publication quality and problem of multiple estimates from a study, here we propose to assign different weights to each meta-observation in our analysis. The proposed weights consist of two components. The first component is the quality index, in which we assign a higher weight (equal to 2) to those primary studies published in good quality journals and a lower weight (equal to 1) to the others. The second one is the importance index, which is given by the number of data sets

---

[10] These possibilities included the non-linear or covariate effects, or the mixed logit models with different distributions on the random parameters.

used divided by the number of estimates published in the primary studies. This index is used to correct the relative importance of each meta-observation, since it could be the case that the more estimates produced from the same data set, the less informative is each of these estimates. The aggregate weights are computed as the product of the quality and importance index. The model in which each observation is weighted by the aggregate weights is estimated by WLS, and the results are presented in Column III of Tables 2.7-2.9.

A natural step forward in the estimation is to combine the two WLS estimators. The results of Column IV in Tables 2.7-2.9 are computed by means of WLS with weights equal to the product of the sample size of the original studies (WLS1, as in Column I) and the quality-importance weights (WLS2, as in Column III). Though we bring some subjectivity into the analysis by introducing the weights in WLS2 and WLS3, our intention here is to compare the results with the previous models and check the robustness of the meta-regressions in the fixed and random effects models.

The meta-regression results will be discussed in the following subsections.

Table 2.7 Results of meta-regression of reliability ratio (RR)

Categories	**Variables**	**Fixed effects (WLS1[a]) I**	**Mixed effects II**	**WLS2[b] III**	**WLS3[c] IV**
	Constant	1.400***	1.286***	1.424***	1.541***
		(5.83)	(9.77)	(7.34)	(3.29)
Data type	*SP*	0.444	-0.126	-0.171	0.171
		(1.60)	(-0.59)	(-0.77)	(0.36)
Unobserved heterogeneity	*HET*	-0.168	-0.107	-0.162	-0.174
		(-0.53)	(-0.46)	(-0.75)	(-0.44)
Utility Specification	*SD*	-1.421***	-0.513**	-0.737**	-1.366***
		(-7.31)	(-2.51)	(-2.53)	(-4.88)
Travel time measurement	*TT_con*	0.198	0.264	0.088	-0.176
		(0.60)	(1.52)	(0.36)	(-0.30)
Reliability measurement	*MM*	0.184	0.792***	0.674***	0.273
		(0.61)	(2.78)	(2.65)	(0.78)
Tau^2 estimate			0.3697		
Adj R-squared		0.7994		0.7737	0.7298
Probability value F-test		0.0000		0.0000	0.0000
Number of observations		74	74	74	74

Note: Significance is indicated by ***, **, and *, referring to significance at the 1%, 5%, and 10% level, respectively, with t-values in parentheses.
[a] WLS weighted by the sample size of the primary studies, which serves as the proxy for the inverse standard errors of the effects sizes; [b] WLS weighted by the assigned quality-importance indexes; [c] WLS weighted by the product of the assigned quality-importance indexes and the sample size of the primary studies.

Table 2.8 Results of meta-regression of schedule delay early ratio (SDER)

Categories	Variables	Fixed effects (WLS1) I	Mixed effects II	WLS2 III	WLS3 IV
	Constant	0.762**	0.638***	0.576***	0.594**
		(2.14)	(2.98)	(3.87)	(2.49)
Data type	*SP*	-0.036	0.013	0.179	0.171
		(-0.11)	(0.08)	(1.62)	(0.82)
Choice type	*BET*	0.310	0.357**	0.504***	0.472***
		(1.42)	(2.01)	(3.57)	(3.06)
Unobserved heterogeneity	*HET*	0.039	0.059	-0.195	-0.166
		(0.19)	(0.37)	(-1.51)	(-1.24)
Trip mode	*PUB*	-0.361***	-0.394***	-0.387***	-0.359***
		(-3.77)	(-4.09)	(-4.07)	(-3.89)
Trip purpose	*COM*	0.114	0.122	0.101	0.107
		(1.16)	(1.21)	(1.06)	(1.16)
Utility Specification	*R*	-0.259**	-0.176	-0.232*	-0.255**
		(-2.21)	(-1.42)	(-1.98)	(-2.28)
Travel time measurement	*TT_con*	-0.263	-0.159	-0.000	-0.043
		(-1.34)	(-1.27)	(-0.00)	(-0.23)
Tau^2 estimate			0.1035		
Adj R-squared		0.8858		0.8863	0.9089
Probability value F-test		0.0000		0.0000	0.0000
Number of observations		69	69	69	69

Note: Significance is indicated by ***, **, and *, referring to significance at the 1%, 5%, and 10% level, respectively, with t-values in parentheses.

Table 2.9 Results of meta-regression of schedule delay late ratio (SDLR)

Categories	Variables	Fixed effects (WLS1) I	Mixed effects II	WLS2 III	WLS3 IV
	Constant	2.865**	1.778**	1.278**	2.378***
		(2.29)	(2.33)	(2.23)	(2.67)
Data type	*SP*	0.098	0.400	1.107***	0.922
		(0.09)	(0.70)	(2.65)	(1.20)
Choice type	*BET*	0.042	0.550	0.201	-0.506
		(0.06)	(0.88)	(0.37)	(-0.86)
Unobserved heterogeneity	*HET*	-1.693**	-1.431**	-1.310**	-1.493***
		(-2.40)	(-2.53)	(-2.59)	(-2.81)
Trip mode	*PUB*	-0.784**	-0.757**	-0.957**	-0.893**
		(-2.37)	(-2.23)	(-2.59)	(-2.48)
Trip purpose	*COM*	0.881**	1.063***	1.262***	0.966***
		(2.63)	(3.02)	(3.43)	(2.75)
Utility Specification	*R*	-2.283***	-1.507***	-1.612***	-2.587***
		(-5.04)	(-3.12)	(-3.66)	(-6.09)
Travel time measurement	*TT_con*	-1.458**	-0.810*	-0.437	-1.126
		(-2.12)	(-1.81)	(-0.96)	(-1.64)
Tau^2 estimate				1.323	
Adj R-squared		0.7956	0.7956		0.7945
Probability value F-test		0.0000		0.0000	0.0000
Number of observations		67	67	67	67

Note: Significance is indicated by ***, **, and *, referring to significance at the 1%, 5%, and 10% level, respectively, with t-values in parentheses.

2.5.1 Data types

The results in Tables 2.7 to 2.9 indicate that the use of the stated preference data has no significant effect on the RR, SDER, and SDLR estimates in our meta-regression, whereas Brownstone and Small (2003) concluded that SP underestimated VOT and VOR substantially. A possible explanation for this phenomenon is that the SP may underestimate both the VOT and the VOR in a systematic but equal-proportional way, and therefore, this downward bias effect is cancelled out by taking the ratio of these two. Similarly, this argument may also hold for the cases of the SDER and SDLR estimates.

2.5.2 Utility specification

Here the utility specification means whether the reliability and scheduling variables are included in estimated model in the primary studies. In the analysis of the RR, the explanatory dummy 'SCHEDULE' denotes the inclusion of scheduling variables in the estimation model. In the analyses of schedule delay ratios, we use the explanatory dummy 'RELIABILITY' to indicate the existence of reliability variables in the same estimation model.

As shown in Tables 2.7 to 2.9, the meta-regression results strongly suggest that including both reliability and schedule variables would have a significant negative effect on RR estimates, as well as on SDER and SDLR estimates. The results are very robust across all the models, in particular for RR and RSDL estimates.

The mis-specification of the choice model may indeed be expected to cause biased estimates for the reliability and scheduling attributes. It is not easy to distinguish between the concepts of reliability and expected scheduling delay, and, statistically, they are positively correlated. Therefore, besides the specification issue of the choice model, it is equally important to have a well-designed choice experiment that is capable of estimating separately the components of reliability and the scheduling variables.

2.5.3 Types of choice set

The meta-regression results in Tables 2.8 and 2.9 show that there is a positive effect on the RSDE estimates obtained from the between-mode choice set; nevertheless, there is no significant difference in the RSDL estimates between the between-mode and the within-mode choice sets. It seems inappropriate to draw any conclusion based on this outcome, since, in principle, if the researchers can correctly model the choice behavior indicated (or observed) by the travelers, the resulting reliability or scheduling estimates should not be too different from each other in different types of choice contexts. However, we are aware that there may exist higher variation in the estimates based on the between-mode choice type questions because of the more complex choice situations faced by the respondents. In such a case, the resulting estimates could be sensitive to the model specification, the possible correlation between alternatives or parameters, and (or) the different nested structure assumed in the logit models.

2.5.4 Trip purposes, modes, and unobserved heterogeneity

The commuting trip is usually considered to have relatively strong scheduling concerns. Thus, it is natural to anticipate some positive effect on the estimates of the scheduling ratios for the commuting trips. The results of the RSDL confirm this, as we can see that all the meta-regression models have significant positive effects on the RSDL for the commuting trips.

For the effect of travel modes, we find that there is a significant negative impact on the scheduling ratio estimates for the public transport mode. One of the possible explanations for this outcome could be that public transport travelers have greater flexibility in allocating their preferred arrival times. In other words, since the public transport users usually have to accommodate the preferred arrival times to the presented timetable, it is possible that these people have less restricted preferred arrival times than the private car users. As a result, the penalty of being early or late is relatively small for public transport compared with that for car.

Regarding the effect of accounting for unobserved heterogeneity, only the results on the RSDL estimate are significant across all models. In fact, the consideration of accounting for unobserved heterogeneity may have different impacts on values of time, reliability, and schedule delay, depending very much on how the researchers specify the choice models. As a result, there will probably be a mixed effect on the ratio of reliability and schedule delay estimates. Whether the consideration of unobserved heterogeneity has certain systematic effects on the empirical estimates requires further information on modeling details and a richer database.

2.5.5 Different measurement attributes

In the investigation of different measurements of travel time, we find out that there is almost no impact on the reliability nor on scheduling ratios if the VOT refers to the congested travel time. This is a rather surprising result, since the general belief is that the value of congested travel time is higher than the value of (mean) travel time, so smaller RRs should be expected in this case. One possible explanation is that the high value of congested travel time obtained in many studies may be because the reliability and scheduling effects were not accounted for (or only implicitly addressed) in these studies. Therefore, once the reliability and scheduling effects are made more explicit in the model, it is likely that the value of congested travel time will be reduced.

Another finding for the measurement impact is that the RR is generally higher when estimated by the definition of ‘Min-Max’. It may be doubted whether we assumed a proper distribution in correcting the value of reliability in Section 2.4.2, although the assumption of a uniform distribution seems more conservative than of a normal distribution. Another possible explanation is that some travelers may consider the ‘Min-Max’ attribute in the choice experiment as still a highly possible outcome for the average delay, treating it as the indicator of a distribution that is in reality wider than the bounds given, and therefore they seem to exhibit more aversion to the definition of Min-Max than to other definitions/ measurements of unreliability.

2.5.6 Implied reliability, schedule delay early and late ratios

The estimated meta-regressions can be used to predict values for the RR, SDER, and SDLR for different modes, purposes, and utility specifications. Table 2.10 provides the values derived from the meta-regression model WLS3. The utility specification ‘Full’ means that both reliability and schedule delay variables are specified in the utility function. It is clear that the effect of the utility specification has a large impact on these ratios. The RR drops from 1.71 to 0.35 when schedule delay variables are included in the utility function. On the other hand, the SDER drops from 0.87 to 0.62 and the SDLR drops from 4.27 to 1.68 in the case of commuting car trips, when the reliability variable is included in the scheduling model. For the implied values of SDER and SDLR, the variations by trip purposes and modes are also not trivial.

Table 2.10 Implied RR, SDER, and SDLR based on meta-regressions of WLS3

Data	Utility	Purpose	Mode	RR	SDER	SDLR
SP	Mean-variance	All	All	1.71	-	-
SP	Full	All	All	0.35	-	-
SP	Scheduling	Commuting	Car	-	0.87	4.27
SP	Scheduling	Commuting	PT	-	0.51	3.37
SP	Scheduling	Others	Car	-	0.77	3.30
SP	Scheduling	Others	PT	-	0.41	2.41
SP	Full	Commuting	Car	-	0.62	1.68

Note: We do not derive the implied SDER and SDLR for the categories of ‘public transport’ and ‘non-commuting car trip’ with the ‘Full’ utility specification. It is because there is no empirical estimate in these categories, and therefore it may not be appropriate to make the ‘out-of-sample’ prediction based on the results of our meta-regressions.

2.6 Conclusions

During the last decade, in the academic literature, reliability and scheduling delay of travel time have been considered as important factors in traveler’s decision making. Many researchers have attempted to model the reliability and scheduling delay attributes into a traveler’s choice model. As a result, a wide range of estimated values has been produced owing to the different data types or methodologies used in the valuation. Our aim in the present study is to analyze the explanatory factors that systematically affect our variables of interest—reliability ratio (RR), scheduling delay early ratio (SDER), and scheduling delay late ratio (SDLR) by means of a multivariate statistical technique: meta-analysis.

To make reliability estimates, evaluated by various measurements, comparable, we began the analysis by adjusting these estimates with some transformation ratios discussed in Section 2.4.2. In the 16 studies we considered, the mean of the adjusted reliability ratios is around 1.33, and the mean values of the schedule delay early ratio and schedule delay late ratio are around 0.75 and 1.65 respectively. We then used several multivariate regression models to further explore the sources of variations between empirical estimations of the RR, SDER, and SDLR. Explanatory variables included in our meta-analysis are: the type of preference data; the choice type; the trip mode; different time and reliability measurements; and the inclusion of schedule and reliability attributes.

We find that, as expected, the inclusion of both reliability and scheduling attributes (SDE, SDL) in the choice model leads to lower estimated values for both the reliability and the schedule delay attributes. Regarding the types of choice data, we do not find any significant difference on the RR, RSDE, and RSDL estimates between stated and revealed preference data. Differences in estimated values may probably be canceled out because we focus on ratios of values here.

Another finding is that the RR estimated by the Min-Max measurement is considerably higher than the one estimated by other measurements, even though we have corrected the RR estimates to be comparable before the meta-analysis. One possible explanation could be that the travelers interpret the 'Min-Max' (uncertain travel time bounds) and the 'unreliability' (some possible travel time realizations) attributes differently in the choice experiment from what is intended by the research.

Our analysis raises the interesting question whether the valuation of reliability or scheduling variables should be based on within-mode choice or between-mode choice type questions. Though the results show that the between-mode choice type has higher estimates in SDER, they do not provide the same evidence in SDLR estimates. Here we can only suspect that between-mode choice type questions may create more variation in empirical valuations as a result of the complex choice situations faced by the respondents.

It remains unclear whether accounting for unobserved heterogeneity has a significant influence on the RR, SDER, and SDLR estimates in our meta-analysis. Nevertheless, we believe that accounting for unobserved behavior heterogeneity, e.g. nested correlations among choice alternatives, more general error structure forms, or unobserved random effects in individuals (randomizing the parameters associated with some attributes), in a more sophisticated manner will result in more accurate estimates, and this is where future research should be heading.

The implied values of the RR, SDER, and SDLR derived in Table 2.10 generalize the main result in our meta-analysis. It provides the insight into how values of the RR, SDER, and SDLR vary in different circumstances. However, because of the limitations of the existing empirical studies in this area, it is not appropriate to predict these ratios for all specific circumstances. It is therefore important that future work should be carried out on reliability and schedule delay estimates in different contexts of trip, and special caution is advised if one wants to apply the values in Table 2.10 in practice.

CHAPTER 3

3 A STUDY OF THE PERCEPTION OF TRAVEL TIME RELIABILITY IN THE SP SURVEY BY USING IN-DEPTH INTERVIEWS

3.1 INTRODUCTION

It is widely accepted that travel time reliability is an important factor in travelers' decision making, and that the benefit of reliability improvement has to be taken into account in current cost-benefit analysis (CBA) frameworks. However, reliability is – unlike factors such as time and costs – not straightforward to specify for a general audience, and therefore the issue of how to present the reliability information effectively and clearly to the respondents becomes crucial in value of reliability (VOR) stated preference (SP) surveys.

Although quite a considerable amount of work has already been done in the area of valuing reliability, there is little consensus about the preferred reliability presentation format (e.g. literal description, clock-face, bar chart, summarized in the Section 3.2 below). Different presentation formats have their own strengths and weaknesses, and may therefore be preferred according to different criteria. Since we were not able to find any systematic knowledge on whether people would prefer, or better understand, certain presentation formats over others, we organized 30 face-to-face interviews to fill this gap in the present knowledge. The objectives of these interviews were to:

- test the respondents' understanding of different reliability presentation formats;
- investigate the respondents' impressions of these presentation formats with regard to several aspects, e.g. clearness, ease of understanding, and visual attractiveness;
- collect the respondents' preferences for the presentation formats with reference to several aspects. In the analyses of these interviews, we paid special attention to the effect of

education[11], to see if all the respondents could understand the questions and especially the presentation of unreliability.

This study was part of an ongoing research project that eventually will lead to a large SP study on the VOR in the Netherlands[12]. The outcomes from the face-to-face interviews will therefore be used to make the recommendations for the content of the formal questionnaire. The remaining sections of this chapter are organized as follows. Section 3.2 reviews the reliability presentation formats that have been used in empirical research. In Section 3.3, the interview questionnaires and the eight formats for the presentation of unreliability are described, together with the respondent recruitment process. In Section 3.4, the outcomes of the tests of the different formats for presenting unreliability to respondents are presented. We give our general impressions and make a final preference for one of the eight presentation formats. Finally, Section 3.5 gives the conclusions from this study. Appendix 3 includes – as an example - the complete eight formats used in the face-to-face interviews for car non-scheduled trips (translated from Dutch).

3.2 OVERVIEW OF EMPIRICAL SP STUDIES IN PRESENTING RELIABILITY

The travel time reliability represents the most difficult attribute in the stated preference experiment as the concept of reliability is difficult to present. Unlike the travel time and travel cost attributes, which usually relate to a specific trip in a stated choice experiment, the reliability often relates to the traveler's experience over multiple trips. As discussed in Chapter 2, the concept of travel time reliability involves two dimensions: a frequency or probability dimension – how often the delay occurs; and a magnitude dimension – how big the delay is when it occurs. The amount of information contained in the reliability attribute is much larger than the information in the other attributes. As a result, the way of presenting travel time reliability in the

---

[11] Other factors, such as gender or age, may also have the influence on how respondents interpret the information of reliability in the SP experiment. Nevertheless, we believe that education plays a very important role on understanding the difficult (statistical) concept of reliability. Given the fact that our interview sample is small, we will only focus on the effect of education here.

[12] The preliminary design project of valuing travel time and reliability was carried out by Significance, VU University Amsterdam, and John Bates for the Netherlands Ministry of Transport Public Works and Water Management, under contract to the AVV Transport Research Centre and the KiM Netherlands Institute for Transport Policy Analysis.

literature varies considerably across studies because of the different thoughts and considerations of the researchers.

To reduce the amount of information presented in the reliability attribute and to facilitate the respondents' task of digesting information on *probability*, many researchers used "5 or 10 possible travel/delay times with *equal chance"* to describe the travel time reliability (see, e.g., Small et al., 1995, 1999; Bates et al., 2001; Hollander, 2005.). Some researchers used literal description to illustrate the reliability attribute, while some thought that graphical presentation can help the respondents to understand the attribute. Some researchers even believed that the respondents are capable of understanding the complicated histogram of travel time distribution if good instruction is provided in the computer-aided personal interviewing (CAPI) program (e.g. Copley et al., 2002). We summarize the reliability presentation formats that have been used in the empirical SP studies in Table 3.1.

Table 3.1 Presentation of travel time reliability in SP studies

No.	**Country of study**	**Author**	**Year**	**Presentation of reliability**	**Visual or literal**
1	US	Small et al.	1995	5 individual travel times, each has an equal chance	Verbal
2	US	Small et al.	1999	5 individual arrival times, each has an equal chance	Verbal
3	US	Small et al.	2005	Frequency of Unexpected Delays of 10 minutes or more: 1 day in 5	Verbal
4	UK	Bates et al.	2001	10 possible delay times is presented graphically in a circular, clock-face format	Graphical
5	NZ	Hensher	2001	Uncertainty time is presented in a vertical bar	Graphical
6	UK	Copley et al.	2002	Graphical presentation of journey time information as a distribution histogram	Graphical
7	UK	Hollander	2005	5 possible arrival times (equal chance) are shown graphically by a series of vertical bars	Graphical
8	US	Bhat and Sardesai	2006	Maximum and usual door-to-door travel time	Verbal
9	US	Tilahun and Levinson	2005	Travel time distribution graphically	Graphical
10	NL	Tseng et al.	2005	Maximum and minimum time in congestion	Verbal
11	NL	Tseng et al.	2006	2 mass points of arrival (delay) times and their associated probabilities	Verbal

3.3 DESCRIPTION OF INTERVIEW QUESTIONNAIRE

3.3.1 The interview setup

The face-to-face interview started with some brief screening questions, in order to be able to assign the interviewee to one of four possible trip types:

1. Car non-scheduled trips (CN): traveling by car as a driver without scheduling consideration;
2. Car scheduled trips (CS): traveling by car as a driver with scheduling consideration;
3. Public transport non-scheduled trips (PN): traveling by public transport without scheduling consideration;
4. Public transport scheduled trips (PS): traveling by public transport with scheduling consideration.

If the timing (departure and/or arrival times) of the trip is important for the traveler, the trip is called a 'scheduled trip', otherwise it is called a 'non-scheduled' trip.

After these sorting questions, the interviewer asked the interviewee to imagine that a certain type of trip was to be considered during the whole interview. The rest of the interview questionnaire is outlined in Table 3.2:

Table 3.2 The structure of the face-to-face interview questionnaire

Section	Content
I	General information on the reliability perception
II	Test questions for the reliability presentation formats
III	General impression of the reliability presentation formats
IV	Preference concerning the reliability presentation formats
V	Background questions

3.3.2 The tested presentation formats

In the selection process of presentation formats to be used, the starting point was the list of reliability SP studies summarized in Table 3.1. Following Hamer et al. (2005), we used a series of possible travel times (or arrival times) to describe travel time unreliability in this study. Thus, we discarded the formats using different types of definition of travel time reliability[13] (e.g, Hensher, 2001; Tseng et al. 2005; Bhat and Sardesai, 2006). In Black and Towriss (1993), (cited from Bates et al., 2001 and Small et al., 1999), it was suggested that people could interpret a 5 mass-points (compared with a 10 mass-points) distribution of travel times reasonably well. We also agreed with this point, and found that a 5 mass-points distribution of travel times is used relatively more widely than others (Small et al., 1995; Small et al., 1999; and Hollander, 2005). Consequently, we adjusted all the tested presentation formats in such a way that each of them can represent a distribution of 5 possible travel times.

We selected eight presentation formats, denoted Formats A to H, to be tested. A brief description of these formats is as follows:

- Format A: a verbal description (without any graph) of 5 possible travel times on 5 different lines. This format is adopted from Small et al. (1999).
- Format B: a clock-face presentation of 5 possible travel times. This presentation format is a variant of the one used by Bates et al. (2001), where there are 10 possible travel times in the clock-face circle.
- Format C: 5 ‘bars’ represent 5 possible travel times. The lower end of the bar gives the departure time and the top end of the bar gives the arrival time. Travel time duration is therefore implied by the length of the bar. This presentation format was used by Hollander (2005).
- Format D: Format D is similar to Format C. The only difference is that the sequence of the 5 possible travel times, which is ordered by the trip time in Format C, is randomized in Format D.
- Format E: histogram of the travel times distributions. The horizontal axis identifies the possible travel times/arrival times, and the vertical axis represents the percentage of trips. This format is used by Copley et al. (2002).

---

[13] Some definitions like ‘uncertainty allowance’ or the ‘spread of maximum and minimum /usual travel time’ are used for the measurement of unreliability. Since the aim of the present face-to-face interviews is to investigate to the best format of presenting a **distribution** of possible travel times, it is therefore not possible to include them in our analysis.

- Format F: Format F is similar to Format E. The only difference between these two is that Format E uses *percentage* (xx%) to represent the likelihood of certain travel time, while Format F uses a relative *frequency* (x out of 5 trips) to represent the same information.
- Format G: Format G is a new format. In this format, we show some possible travel time/arrival time intervals and the associated frequencies to the respondents. The consideration here is that these possible arrival time intervals are closer to reality, and may be more representative for the concept of unreliability.
- Format H: Format H is a combined version of Format A and G. The 5 possible travel time/arrival time intervals are described on 5 different lines.

The eight presentation formats are given in Appendix 3A for the car non-scheduled (CN) trips.

3.3.3 Recruitment and interview process

The 30 interviewees were recruited from the personal network of relatives and friends of the Department of Spatial Economics, Vrije Universiteit Amsterdam. To enhance the credibility of the interview outcomes, we paid special attention to selecting representative interviewees from every education level, as well as with respect to the types of trips (CS, CN, PS, and PN) they made. All the interviews were done during October and November 2006. An interview usually took 40 – 60 minutes, while a few took more than 1 hour. The respondent was given 15 euros as a reward for participating.

3.4 RESULTS FROM THE FACT-TO-FACE INTERVIEWS

The number of interviews was 30 in total, with 8 respondents in the CN, 10 respondents in the CS, 3 respondents in the PN, and 9 respondents in the PS versions. In the following subsections, we will present the results according to the sequence of questions in the questionnaire. Frequency tables are presented for each question. In particular, we will focus on the group of lower-educated people[14].

[14] These are respondents with education levels no higher than LO, MAVO/VBO/VMBO/LBO, and MBO.

3.4.1 General information on the reliability perception

Before giving any description or explanation of unreliability (uncertainty) of travel time to the interviewees, we asked the respondents how they would describe the unreliability/uncertainty of travel time, and how they think of uncertainty in practice.

Table 3.3 QI_1. Can you indicate which factors according to you, in practice, would account for unreliability of travel time? (you can choose more than one option)

	Lower education		Higher education		All respondents	
	Freq.	%	Freq.	%	Freq.	%
Average travel time	11	61.1	10	28.6	21	39.6
Maximum travel time	4	22.2	7	20.0	11	20.8
Minimum travel time	1	5.6	8	22.9	9	17.0
Probability	2	11.1	9	25.7	11	20.8
Don't know	0	0.0	0	0.0	0	0.0
Other	0	0.0	1	2.9	1	1.9
Total	18	100.0	35	100.0	53	100.0

After explaining that travel times are usually uncertain in practice, we first asked which elements of travel time people consider when travel time is uncertain. As expected, ‘average travel time’ got more votes than the other elements. The elements of maximum/minimum travel time and probability accounted for around an equal share of the choices. In principle, average travel time, maximum and minimum travel times, and probability are the four most important factors that the travelers would think of to describe the travel time uncertainty in practice. It is interesting to note that higher-educated people tend to think more of ‘probability’ than lower educated people.

Next, the interviewers gave a brief description and explanation for the uncertainty of travel time. In particular, a throwing dice example was given to the respondents to help them to think about the unpredictability of travel times and the probability associated with the realization of a certain travel time event in the context of a mass-point distribution. Questions QI_2 to QI_4 were then asked after the description and the throwing dice example. Because it is important that people are aware of the ‘true’ nature of uncertainty when looking at 5 different travel times, it seems useful to include a text like this in the final SP questionnaire.

Table 3.4 QI_2. How complicated do you find the description of uncertainty of travel times in the text (throwing dice example)?

	Lower education		Higher education		All respondents	
	Freq.	%	Freq.	%	Freq.	%
Not at all complicated	2	16.7	4	22.2	6	20.0
A bit complicated	6	50.0	9	50.0	15	50.0
Reasonably complicated	1	8.3	2	11.1	3	10.0
Considerably complicated	2	16.7	2	11.1	4	13.3
Very complicated	1	8.3	1	5.6	2	6.7
Total	12	100.0	18	100.0	30	100.0

Table 3.5 QI_3. How helpful do you find the example of throwing a dice in the text?

	Lower education		Higher education		All respondents	
	Freq.	%	Freq.	%	Freq.	%
Not at all helpful	2	16.7	5	27.8	7	23.3
A bit helpful	2	16.7	4	22.2	6	20.0
Reasonably helpful	2	16.7	4	22.2	6	20.0
Considerably helpful	4	33.3	2	11.1	6	20.0
Very helpful	2	16.7	3	16.7	5	16.7
Total	12	100.0	18	100.0	30	100.0

Table 3.6 QI_4. To what extent does the description of uncertainty of travel times in the text correspond with how you think of uncertain travel times in practice?

	Lower education		Higher education		All respondents	
	Freq.	%	Freq.	%	Freq.	%
Not at all	2	16,7	6	33,3	8	26,7
A bit	2	16,7	4	22,2	6	20,0
Reasonably	6	50,0	3	16,7	9	30,0
Considerably	1	8,3	2	11,1	3	10,0
Very much	0	0,0	1	5,6	1	3,3
Not applicable, never thought about uncertainty	1	8,3	2	11,1	3	10,0
Total	12	100,0	18	100,0	30	100,0

We can conclude that most respondents find it at least “a bit” helpful, and more than half respondents at least “reasonably” helpful to draw the parallel with throwing dice (Table 3.6). It matches at least “reasonably” for nearly half the respondents who think the description of uncertainty in travel times in the text corresponds with how they think of about uncertain travel times in practice (Table 3.7). For those respondents for whom this is not true, it may be even more important to use the example of throwing dice, because we want them to think about unreliability as it is specified in the SP study.

3.4.2 Test questions for reliability formats

The main objective of this face-to-face interview was to test which reliability presentation format is best understood by the respondents, and to see how the respondents interpreted these presentations. After showing each reliability presentation format to the respondents, in the interview Section II, the interviewer asked some ‘test questions’, associated with that particular format. These test questions were designed to check to what extent the respondents have the “correct” perception of reliability, i.e. the same as intended by the researchers, for each format. To prevent people getting too familiar with the numerical example and learning the answer to previous formats, we varied the attributes (time, cost, and reliability) levels across these eight presentation formats. Furthermore, we also randomized the order of showing the formats, in order to reduce the possible biases from learning or fatigue effects. In other words, Formats A to H were shown in different orders to different respondents. The results of the test questions are now summarized for scheduled and non-scheduled trip versions separately.

Non-scheduled trips

There are two relevant test questions for the non-scheduled trip interviewees (versions of CN+PN=11 respondents). We calculate the percentages of correct, incorrect, and non-response answers for each format. Tables 3.7 and 3.8 summarize the results of these two test questions.

Though the sample size is rather small, it is encouraging to see that the percentages of correct answers for lower-educated people are satisfactory, and in many cases they are even higher than the ones for higher-educated people.

Table 3.7 QII_3. The travel time for trip 1 is more uncertain[a] than trip 2

Format	A	B	C	D	E	F	G	H
Lower education								
% correct	100	100	100	80	80	100	100	100
% incorrect	0	0	0	20	20	0	0	0
% non-response	0	0	0	0	0	0	0	0
Higher education								
% correct	83	67	67	83	67	83	50	83
% incorrect	0	33	17	0	17	17	33	0
% non-response	17	0	16	17	16	0	17	17
All respondents								
% correct	91	82	82	82	73	91	73	91
% incorrect	0	18	9	9	18	9	18	0
% non-response	9	0	9	9	9	0	9	9

Note: [a.]Here, 'more uncertain' can be interpreted either by a larger standard deviation of travel time or a wider range between the maximum and minimum travel times (amount of uncertain time). In this interview experiment, we deliberately assigned the reliability level (a series of 5 possible travel times) in such a way that a more unreliable alternative has a larger standard deviation, as well as a larger uncertain time (difference between the maximum and minimum travel times). From the academic point of view, it may be interesting to see whether people would take the standard deviation of travel time or the amount of uncertainty time into account in the choice decision. However, this point is less relevant for the interview purpose. To eliminate possible confusion, we used the strategy as explained.

Table 3.8 QII_4. With trip 1, I arrive, on average, earlier than with trip 2

Format	A	B	C	D	E	F	G	H
Lower education								
% correct	80	100	-	-	80	60	100	100
% incorrect	20	0	-	-	20	40	0	0
% non-response	0	0	-	-	0	0	0	0
Higher education								
% correct	83	83	-	-	83	67	83	83
% incorrect	0		-	-	0	33	0	0
% non-response	17	0	-	-	17	0	17	17
All respondents								
% correct	82	91	-	-	82	64	91	91
% incorrect	9	9	-	-	9	36	0	0
% non-response	9	0	-	-	9	0	9	9

Note: Since the two alternatives in Format C and D have the same average travel time, there is no correct answer for these two formats.

Scheduled trips

There are four test questions for the interviews concerning scheduled trips (versions of CS+PS=19 respondents). The results are summarized in Tables 3.9 to 3.12.

Table 3.9 QII_1. With trip A, I have a greater probability of arriving earlier than I want compared with trip 2

Format	A	B	C	D	E	F	G	H
Lower education								
% correct	86	71	100	71	71	71	-	-
% incorrect	14	29	0	29	29	29	-	-
% non-response	0	0	0	0	0	0	-	-
Higher education								
% correct	100	50	92	83	83	83	-	-
% incorrect	0	50	8	17	17	17	-	-
% non-response	0	0	0	0	0	0	-	-
All respondents								
% correct	95	58	94	79	79	79	-	-
% incorrect	5	42	7	21	21	21	-	-
% non-response	0	0	0	0	0	0	-	-

Table 3.10 QII_2. With trip A, I have a greater probability of arriving later than I want compared with trip 2

Format	A	B	C	D	E	F	G	H
Lower education								
% correct	-	86	100	71	71	57	-	-
% incorrect	-	14	0	29	29	43	-	-
% non-response	-	0	0	0	0	0	-	-
Higher education								
% correct	-	58	92	92	75	58	-	-
% incorrect	-	42	8	8	25	42	-	-
% non-response	-	0	0	0	0	0	-	-
All respondents								
% correct	-	68	95	84	74	58	-	-
% incorrect	-	32	5	16	26	42	-	-
% non-response	-	0	0	0	0	0	-	-

Table 3.11 QII_3. The travel time for trip 1 is more uncertain than it is for trip 2

Format	A	B	C	D	E	F	G	H
Lower education								
% correct	57	43	29	43	29	43	43	71
% incorrect	29	43	71	29	57	57	29	15
% non-response	14	14	0	28	14	0	28	14
Higher education								
% correct	100	67	50	83	67	83	50	83
% incorrect	0	33	50	17	33	17	50	0
% non-response	0	0	0	0	0	0	0	17
All respondents								
% correct	84	58	42	68	53	68	47	79
% incorrect	11	37	58	21	42	32	42	5
% non-response	5	5	0	11	5	0	11	16

Table 3.12 QII_4. With trip 1, I arrive, on average, earlier compared with trip 2

Format	A	B	C	D	E	F	G	H
Lower education								
% correct	86	86	100	71	86	86	-	-
% incorrect	14	14	0	29	14	14	-	-
% non-response	0	0	0	0	0	0	-	-
Higher education								
% correct	83	50	83	92	83	92	-	-
% incorrect	17	50	17	8	17	8	-	-
% non-response	0	0	0	0	0	0	-	-
All respondents								
% correct	84	63	89	84	84	89	-	-
% incorrect	16	37	11	16	16	11	-	-
% non-response	0	0	0	0	0	0	-	-

The percentages of correct answers to test question QII_3 (Table 3.12) are generally lower than those to the other question, especially for the lower-educated people. Since the non-response percentages are also higher for this question, it is possible that these respondents did not know how exactly to interpret ‘uncertainty’, and they just gave up in answering this question. Another possible reason for the lower percentages of correct answers to this question is that the 5 possible travel times are shown together with 5 resulting possible arrival times in the format, so people may be inclined to link the uncertainty to the level of lateness rather than to the level of unreliability.

The percentages of correct answers for Format B are relatively lower than those for the other formats for these test questions. As for the level of education, we do not find any strong evidence that lower-educated people have lower percentages of correct answers.

3.4.3 General impression of reliability presentation formats

In interview Section III, we asked the interviewees to indicate how they perceive the levels of clearness or difficulty in each presentation format. This is described by five different indicators:

- Clearness of the presentation of reliability
- Ease of making a choice between two alternatives/trips
- Ease of considering all information/attributes
- Attractiveness of the visual appearance

– Ease of answering the test questions in Section 3.4.2.

To make the answers more comparable between formats, we assign scores to different levels of understanding (or difficulty). Thus, we can rank these eight presentation formats according to the average scores they get in the questions. Here we use 1 for 'very unclear', 2 for 'unclear', 3 for 'not very clear', 4 for 'clear', and 5 for 'very clear'. A format with a higher score will therefore be more preferred.

Figures 3.1 to 3.3 below, summarize the average scores (the average of the scores of 30 respondents) in these eight formats based on these five indicators for lower-educated, higher-educated, and total respondents, respectively.

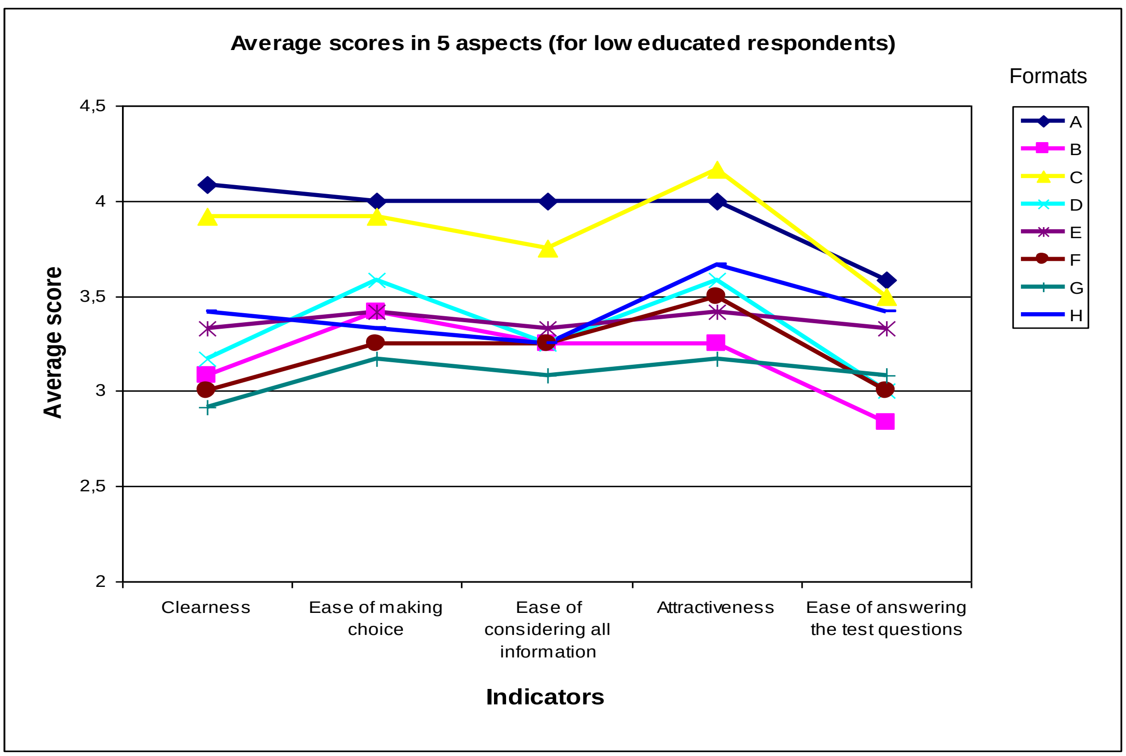


Figure 3.1 The average scores of 5 indicators for Formats A to H: lower-educated respondents

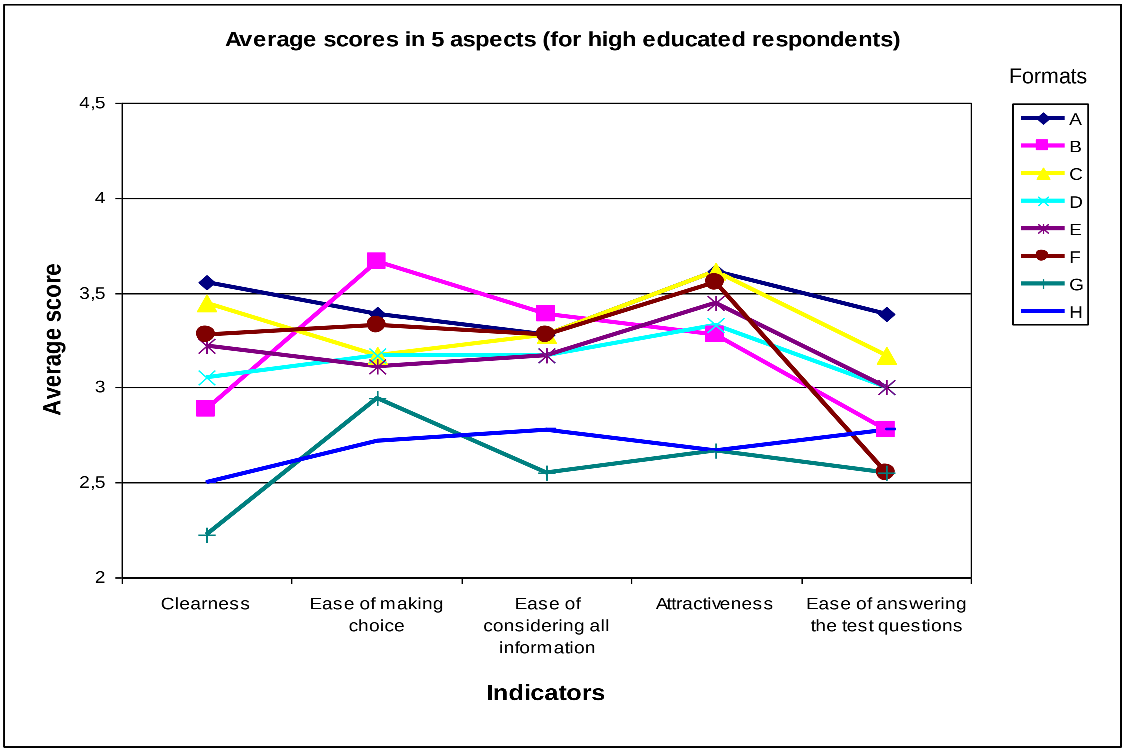


Figure 3.2 The average scores of 5 indicators for Formats A to H: higher-educated respondents

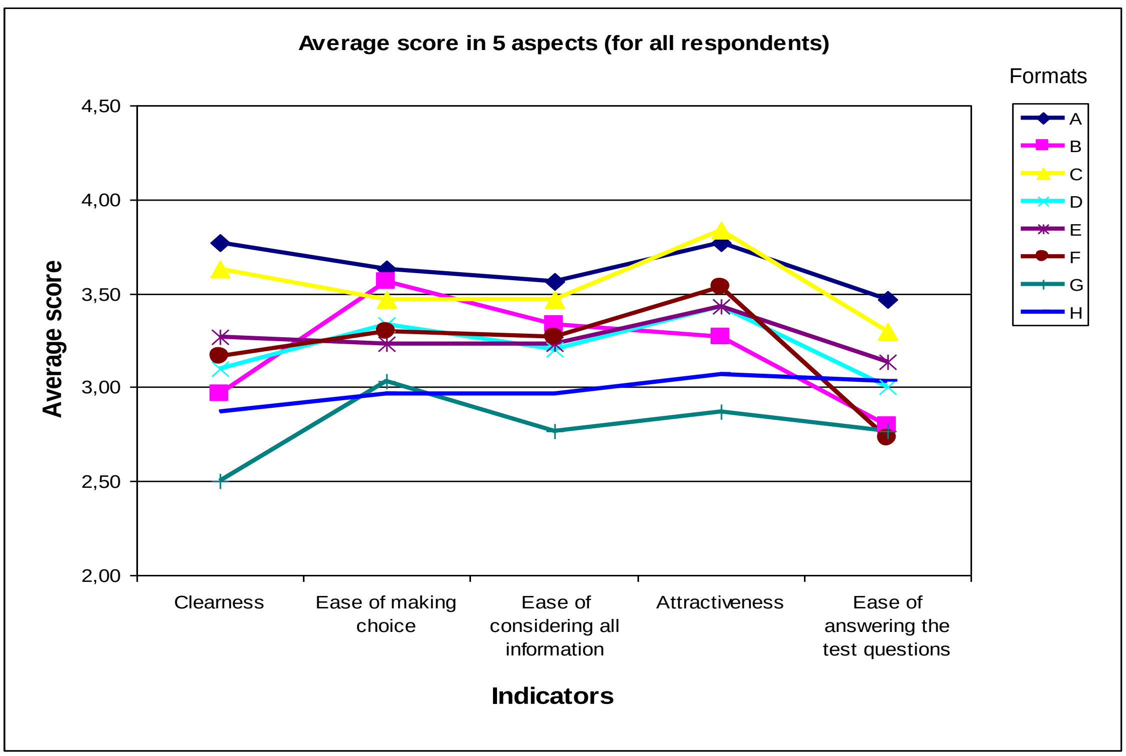


Figure 3.3 The average scores of 5 indicators for Formats A to H: total respondents

In general, Format A is ranked first in most cases. The next is Format C, which also performs well in terms of the scores, and the score differences between Format A and C are small in many cases. It is obvious that Format G and H are among the worst for most of the cases. The reliability information given in these two formats is more complex (showing the travel time intervals rather than some certain travel times), and the respondents seem to have much more difficulty in reading these two formats.

3.4.4 Preference of presentation formats

In interview Section IV, we showed all eight presentation formats once more to the interviewees. The interviewers placed these eight formats on the desk (randomly), and then asked the interviewees to indicate once again their preferences for the format, according to the five indicators discussed in Section 3.3.3 (clearness of the presentation; ease of making a choice between two alternatives; ease of considering all information in the alternatives; visual appearance attractiveness; ease of answering the test questions).

The respondents indicated both the most-preferred and the least-preferred format according to these five aspects. The results of the choice frequency are shown graphically in Figures 3.4 to 3.8. Here we combine the most and the least-preferred format into one graph. Thus, the positive part of the vertical axis represents the frequency that the format is chosen as favorite format; while the negative part represents the frequency that it is chosen as the least preferred format.

From these results we see that the preferences of these respondents are quite diverse. None of the formats is favored (most) by more than half of the respondents in every situation (clear/easy/attractive…etc.). Nevertheless, the simplest Format A performs well consistently, compared with the other formats, and the scores are relatively high on criteria that are more likely to directly affect the quality of responses for repetitive SP questions (ease of making choice, ease of considering all information, ease of answering test questions).

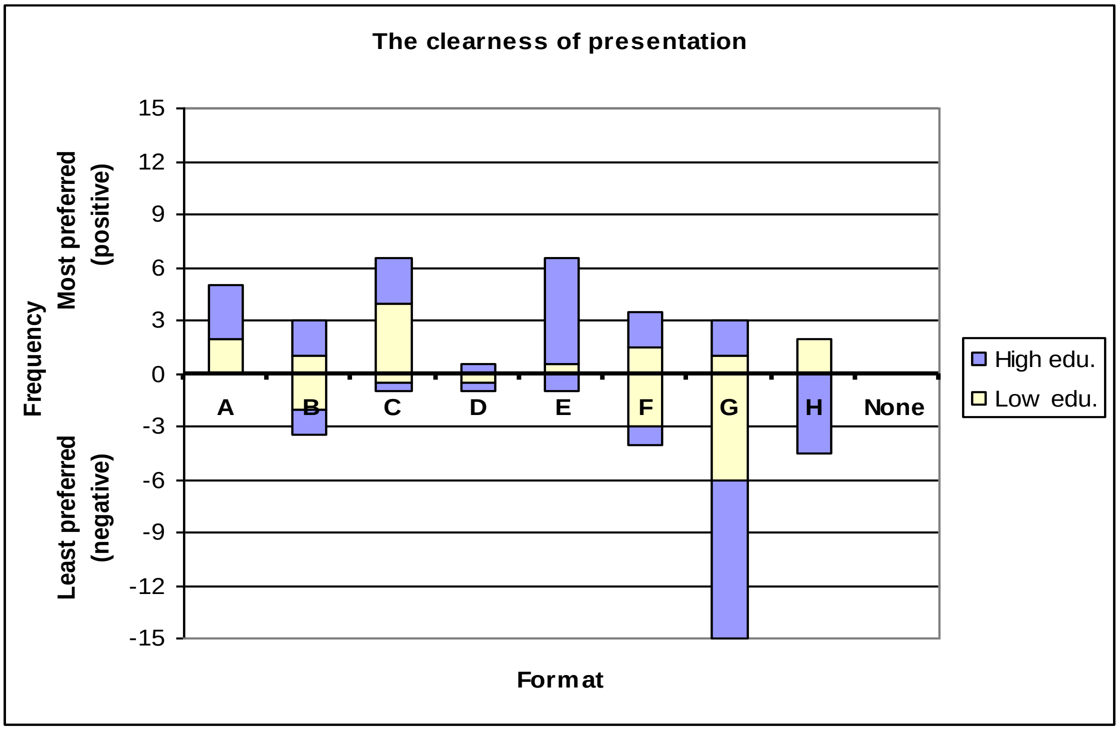


Figure 3.4 The frequency of measuring the clearness of the format

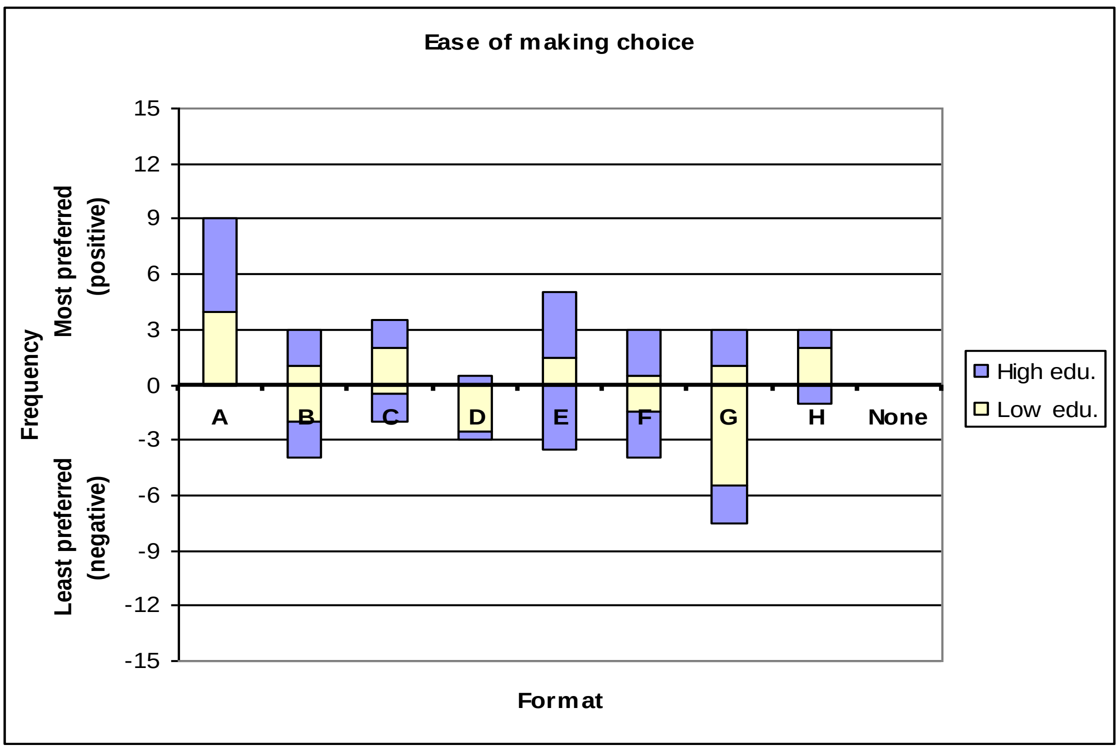


Figure 3.5 The frequency of measuring the ease of understanding the format in making a choice

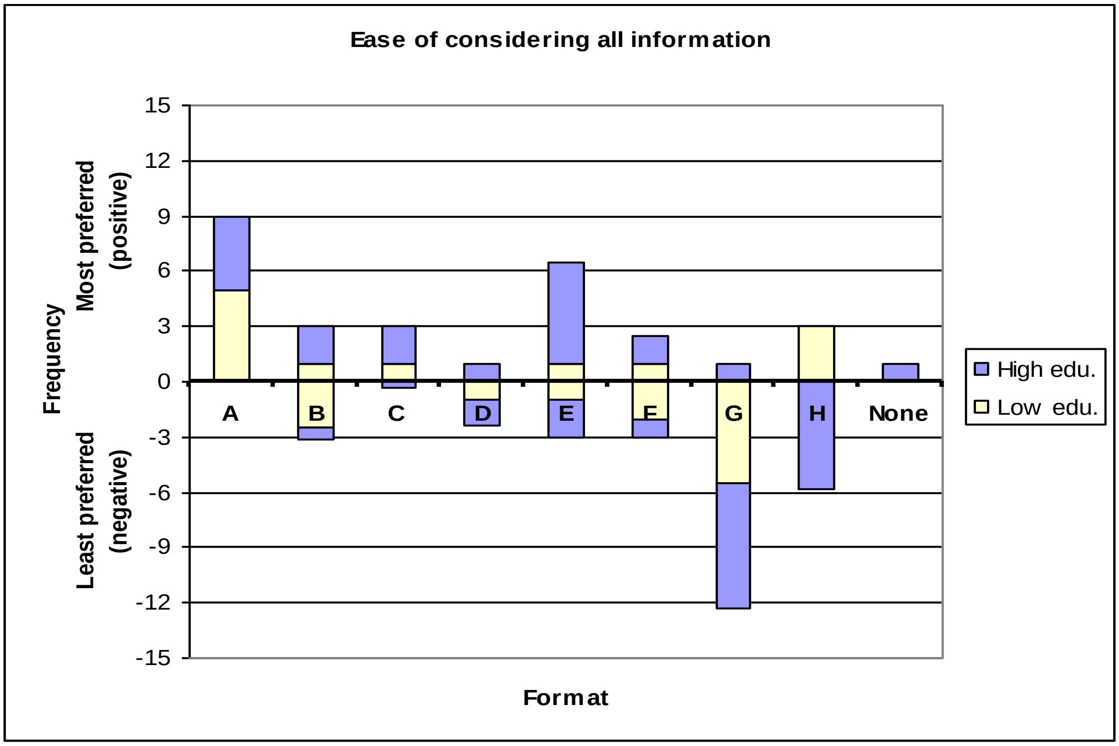


Figure 3.6 The frequency of measuring the ease of understanding the format in considering all information

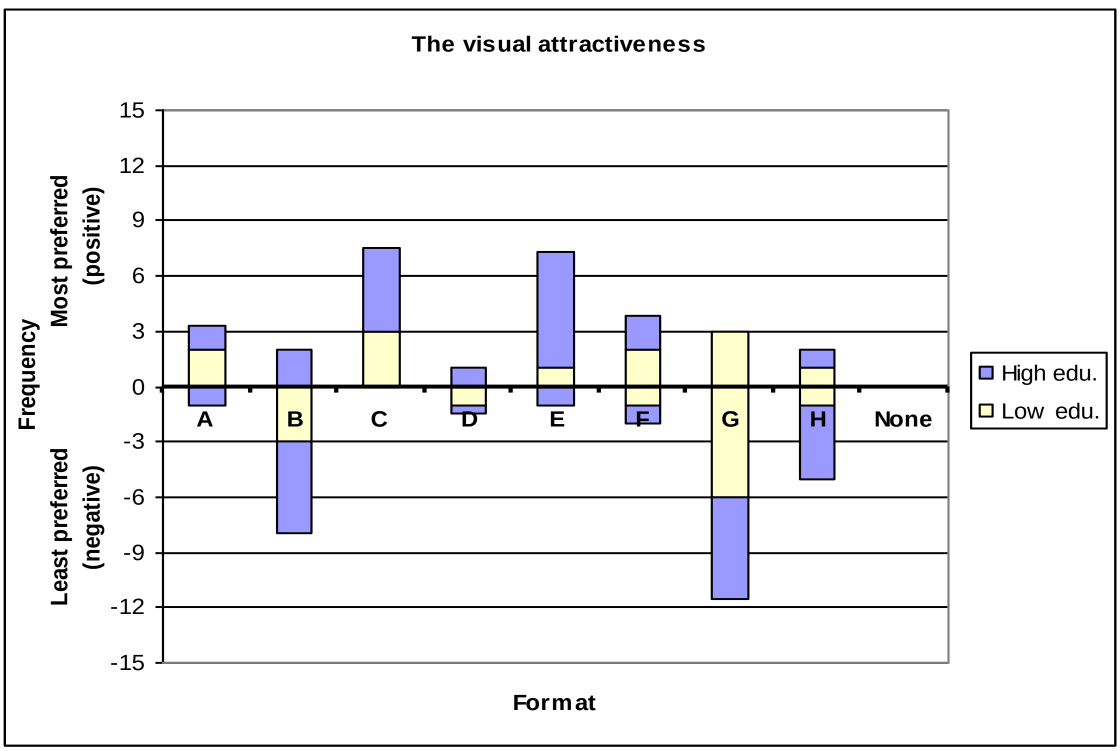


Figure 3.7 The frequency of measuring the format concerning its the visual attractiveness

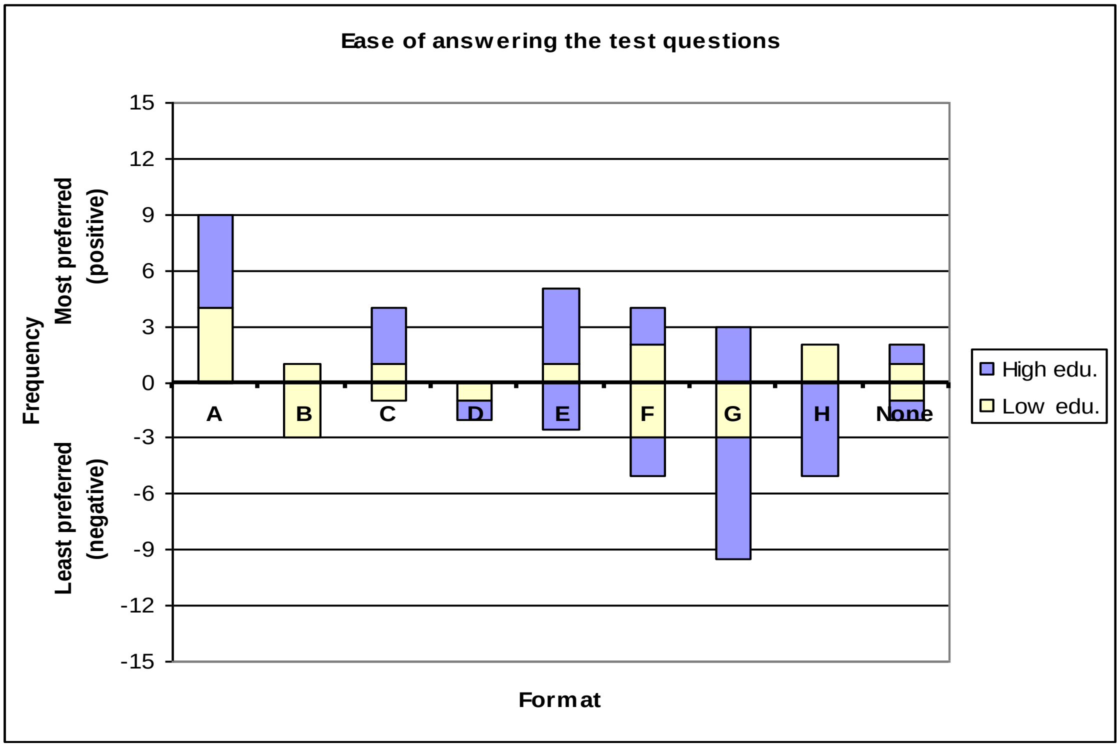


Figure 3.8 The frequency of measuring the format concerning the ease of answering the test questions

3.5 CONCLUSION

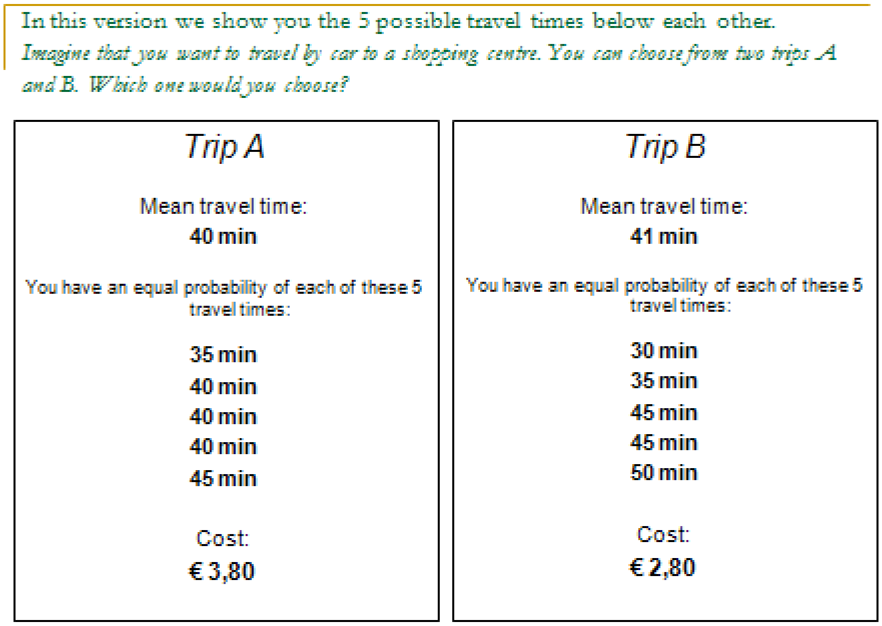
In the face-to-face interviews discussed in this chapter, we first collected information about respondents' perception of travel time unreliability. We then examined several reliability presentation formats that have been used in previous empirical SP studies. To have a thorough assessment of these presentation formats, we not only asked questions about the respondents' subjective preferences, but also asked questions that can be tested objectively to see if these respondents perceived the unreliability as expected. The interview results were analyzed separately for lower- and higher-educated people.

In conclusion, we recommend using the verbal description, Format A, to represent the travel time reliability in the VOR SP experiment. In the interviews questionnaires, Format A was favored by most (a relatively high proportion) of the respondents. Furthermore, respondents' preference for Format A is rather consistent between lower- and higher-educated people. Histograms with unequal probabilities (Formats E and F) seem less promising, since we find that some people have difficulty in reading the probability from the graph, especially lower-educated people. It is also interesting to note that the preferences between Format C and D are quite varied in some cases. Since Format C and D are quite similar in the way of presenting reliability information, we would expect that the preferences of these two formats should be comparable to a certain extent. For this reason, though Format C is ranked second in many cases, we would not recommend this format highly here, and feel second thoughts will be needed if Format C is to be considered seriously.

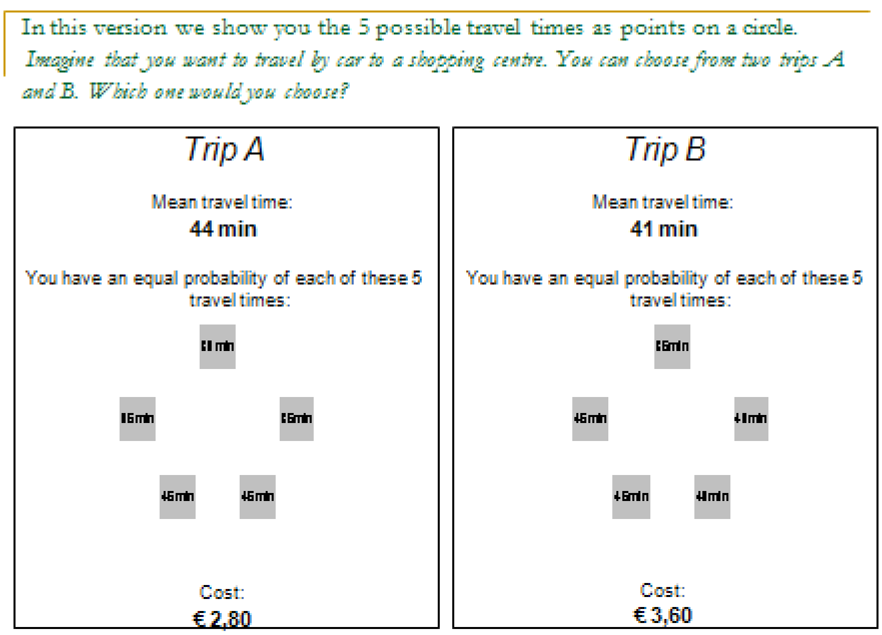
One remark about these interviews is that the test formats are conditional on the 5 possible travel times with equal probability. Thus, we shall emphasize that the above conclusion is contingent on this particular situation and may not apply to all cases. In some cases, when very small differences in probabilities (far below say 20%) in one of the tails of the travel time distribution (often the right tail) are likely to be decisive for choice behavior, the presentation format suggested in this study may no longer be adequate. A trip to the airport is such an example. Because the consequence of missing the flight is usually very severe, even a small probability may play an important role in a traveler's decision making. In such a case, the presentation format E could be an alternative; nevertheless, some proper training in reading the histogram may be required and would be helpful for the respondents.

In the following chapters, we will present the empirical estimates of VOR based on the SP experiments developed for car commuters (Chapter 4) and railway passengers (Chapter 6). Although the face-to-face interviews in this chapter were carried out later than these SP studies, the results from this chapter are consistent with the presentation approach taken in these SP studies. That is, both the SP experiments in later chapters used verbal descriptions rather than graphs or histograms to represent the attribute of travel time reliability.

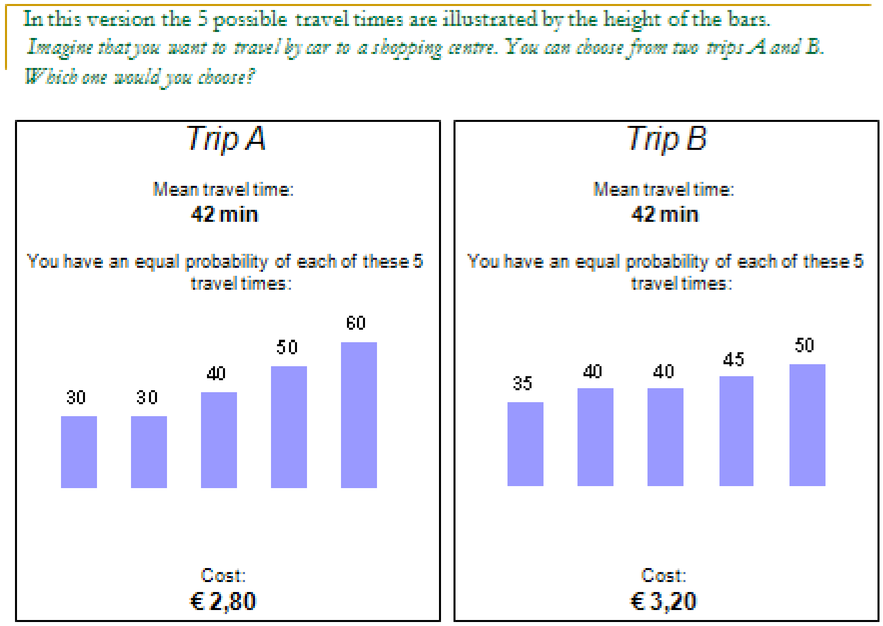
Appendix 3A: Reliability presentation formats for car non-scheduled trips



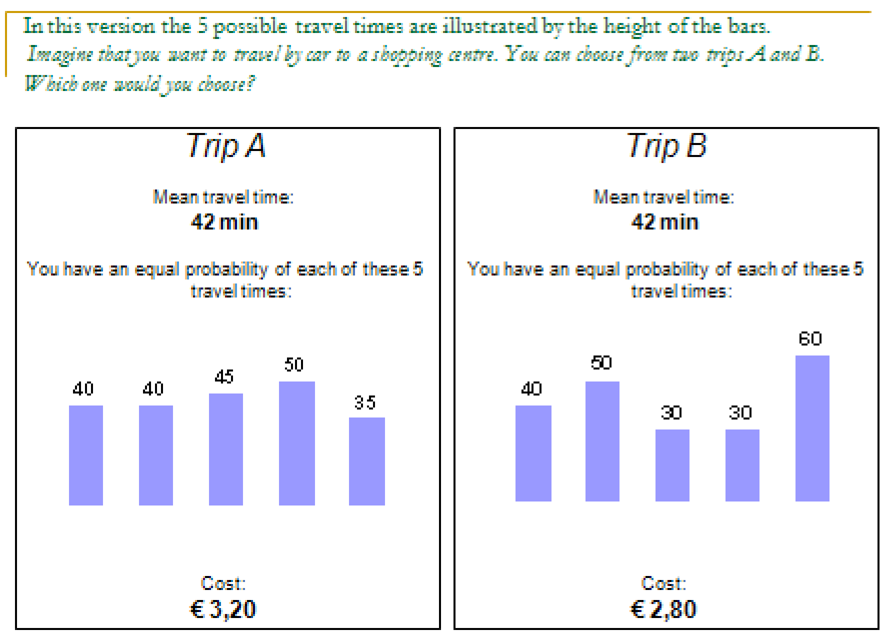
A



B


In this version the 5 possible travel times are illustrated by the height of the bars.
Imagine that you want to travel by car to a shopping centre. You can choose from two trips A and B.
Which one would you choose?
Trip A
Mean travel time:
42 min
You have an equal probability of each of these 5 travel times:
30
30
40
50
60
Cost:
€ 2,80
Trip B
Mean travel time:
42 min
You have an equal probability of each of these 5 travel times:
35
40
40
45
50
Cost:
€ 3,20

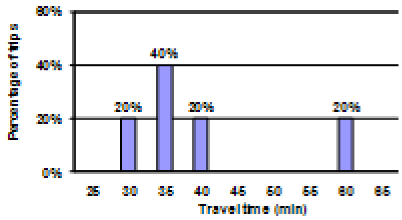
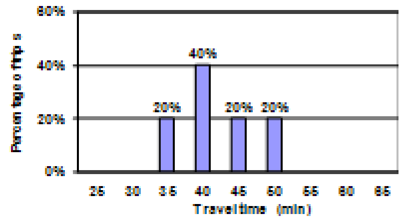
C


In this version the 5 possible travel times are illustrated by the height of the bars.
Imagine that you want to travel by car to a shopping centre. You can choose from two trips A and B.
Which one would you choose?
Trip A
Mean travel time:
42 min
You have an equal probability of each of these 5 travel times:
40
40
45
50
35
Cost:
€ 3,20
Trip B
Mean travel time:
42 min
You have an equal probability of each of these 5 travel times:
40
50
30
30
60
Cost:
€ 2,80

D

In this version the 5 possible travel times are illustrated by the height of the bars. (probabilities as percentage).

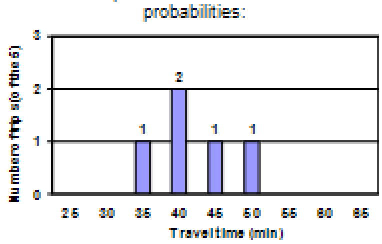
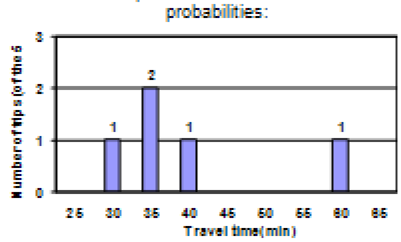
Imagine that you want to travel by car to a shopping centre. You can choose from two trips A and B. Which one would you choose?

Trip A	Trip B
Mean travel time: **40 min**	Mean travel time: **42 min**
These are the possible travel times and their probabilities:	These are the possible travel times and their probabilities:
	
Cost: **€ 3,00**	Cost: **€ 3,40**

E

In this version the 5 possible travel times are illustrated by the height of the bars. (probabilities as number of times out of 5)).

Imagine that you want to travel by car to a shopping centre. You can choose from two trips A and B. Which one would you choose?

Trip A	Trip B
Mean travel time: **42 min**	Mean travel time: **40 min**
These are the possible travel times and their probabilities:	These are the possible travel times and their probabilities:
	
Cost: **€ 3,40**	Cost: **€ 3,00**

F

In this version we show in words and along a time axis how large the probabilities of certain travel times are (as number of trips of the five). The travel times are, thus, not precise, but within certain limits

Imagine that you want to travel by car to a shopping centre. You can choose from two trips A and B. Which one would you choose?

Trip A

Usual travel time:
50 min

These are the possible travel times and their probabilities:

Travel time	Probability
30 – 40 min	1 of the 5 trips
40 – 60 min	3 of the 5 trips
More than 60 min	1 of the 5 trips

Cost:
€ 2,40

Trip B

Usual travel time:
40 min

These are the possible travel times and their probabilities:

Travel time	Probability
30 – 35 min	1 of the 5 trips
35 – 45 min	3 of the 5 trips
More than 45 min	1 of the 5 trips

Cost:
€ 3,60

G

In this version we show in words how large the probabilities of certain travel times are (as number of trips of the five). The travel times are, thus, not precise, but within certain limits

Imagine that you want to travel by car to a shopping centre. You can choose from two trips A and B. Which one would you choose?

Trip A

Usual travel time:
40 min

You have an equal probability of each of these 5 travel times:

Travel time

30 – 35 min
35 – 40 min
40 – 45 min
45 – 50 min
More than 50 min

Cost:
€ 3,40

Trip B

Usual travel time:
45 min

You have an equal probability of each of these 5 travel times:

Travel time

30 – 40 min
40 – 45 min
45 – 50 min
50 – 65 min
More than 65 min

Cost:
€ 2,80

H

CHAPTER 4

4 VALUING TRAVEL TIME, SCHEDULE DELAY, AND UNCERTAINTY: ESTIMATIONS OF A STATED CHOICE EXPERIMENT RELATING TO DUTCH COMMUTERS

4.1 INTRODUCTION

Traditionally, the value of travel time is thought to be one of the largest cost components in cost-benefit analysis of transportation projects, and the reduction of travel time is usually regarded as the main source of benefits that travelers receive from the improvement of a transport facility. However, when the seriousness of road congestion rises considerably, the reliability of travel time may be considered as more important than the savings of travel time for the travelers, particularly when travelers have a schedule constraint. Several reliability-related components, such as the standard deviation of travel time, have therefore been considered in the literature of mode or route choice. Numerous studies have shown the importance of these reliability factors in traveler's choice behavior, and in some cases reliability has acquired an even higher monetary value than the value of travel time savings (see, e.g., Small et al., 1999; Lam and Small, 2001).

Research has generated an enormous literature on empirical estimates of the VOT, but a much smaller one on value of reliability (VOR). One of the main reasons is that it is difficult to measure unreliability in actual situations, so that a common definition of reliability is still lacking. Furthermore, the modeling approaches (utility specifications) of traveler's responses to changes in reliability differ over studies. Our research contributes by providing useful new insights into relatively unknown concepts, such as the value of scheduling costs (at least for the Dutch situation), by updating the VOT for an important target group (i.e. commuters), and by providing insight into the value of uncertainty.

In this chapter, we present the outcomes from a stated choice experiment (SCE) relating to Dutch car commuters who experience congestion on a regular basis. The respondents were offered two different types of stated choice questions. The first four choice sets (part A) contained simple choice structures, in which one single attribute and the cost term varied, thus enabling us to estimate interval values for time, schedule delay (late and early), and uncertainty (or reliability) for each individual. The second set of choice alternatives (part B) included more attributes, which were varied in a systematic way, and more choices that had to be made by the respondents (11 screens were shown). The data from both types of questions are different in nature and hence require a different type of analysis. We start with the more simple analysis of the interval estimates resulting from the first four choice sets. Next, we analyze the second set of choice data by estimating different types of models. This allows us to compare estimates from both approaches.

The aim of this study is twofold. First, it is useful to derive (up-to-date) estimates of important transport concepts, such as the VOTs and reliability, for this particular group of car drivers who experience congestion. Secondly, we study the impact of the individual characteristics on these estimates. By allowing those trait variables to interact with travel time and reliability-related attributes, we are able to obtain the estimates for different groups of travelers.

The remainder of this chapter is organized as follows. Section 4.2 describes the choice experiment and the data used in our empirical assessment. Section 4.3 presents the results from the first part of the SCE, the interval estimates. It includes a short statistical analysis in order to search for factors that explain variation over individuals. Section 4.4 provides some background by discussing the theoretical framework of discrete choice analysis. The estimated models will then be presented in Section 4.5. Finally, Section 4.6 concludes.

4.2 DATA SOURCES AND SURVEY DESCRIPTION

4.2.1 Data collection

The data have been obtained by conducting an (interactive) computer-based survey among Dutch commuters. The questionnaire can roughly be divided into three parts. First, we asked for some socioeconomic characteristics concerning the respondent (such as education and income). Next, the two SCEs were presented to the respondents. And, finally, we asked for their opinion (on issues such as acceptance and effectiveness) about different types of road pricing measures. This chapter presents the data analysis of the SCEs. See Ubbels (2006) for further analyses of the other parts of the questionnaire.

The data collection was carried out by a specialized firm (NIPO) that has a panel of over 50,000 respondents. Since the survey was aimed at respondents who use a car for their commuting journey and also experience congestion on a regular basis, we selected working respondents who drive to work by car two or more times per week, and who experience congestion of 10 minutes or more at least two times a week. An initial analysis revealed that a random sample would result in a relatively low number of women and lower-income groups. In order to allow an investigation of the role of income, it was decided to 'over sample' the lower-income groups and create an equal number of respondents over the various income classes. The final sample includes 1115 respondents. The data were collected during three weeks in June 2004.

4.2.2 Survey

As previously explained, the survey started with some general questions asking for important explanatory variables concerning the characteristics of the respondent (such as income, gender and education). This provided us with a profile of the Dutch commuter who experiences congestion. The respondent was asked to distribute 10 (commuting) trips amongst four alternatives in the choice experiment. The attribute levels of the alternatives were based on answers of respondents about their current travel behavior. In total, the respondents were confronted with 15 (4+11) screens in the experiment: 4 for part A, and 11 for part B.

The first four screens in part A were simpler versions of the choice experiment, in which we only varied the road pricing fee and one other attribute (travel time, shift to earlier arrival time, shift to later arrival time, or uncertainty in travel time) between alternatives. These choice questions were designed in such a way that we can infer interval estimates for individuals' values of time (VOT), values of schedule delay early (VSDE), values of schedule delay late (VSDL), and values of uncertainty (or reliability, VUNC) from the allocation of 10 trips over four alternatives. The units for all parameters values are €/hour. The choice for a particular trip implies an interval in which the VOT, VSDE, VSDL or VUNC must lie for that particular choice situation. The more the ten trips are allocated to a particular alternative, the less dispersed that individual's VOT, VSDE, VSDL or VUNC apparently is. The design of the alternatives developed to derive the VOT estimate is then as follows.

The VOT for car commuting trips in the Netherlands, as used by the Dutch government for official policy evaluations in 2004, was about 8.3 €/hour (see AVV, 2006). Given this value, we defined the following four intervals:

1. € 0 – 4
2. € 4 – 8
3. € 8 – 12
4. > € 12

In order to allocate responses to the above categories, a choice was offered to the respondent, as presented in Table 4.1.

Table 4.1 Design of the first screen: four alternatives to estimate an individual's VOT

	A (group 4)	B (group 3)	C (group 2)	D (group 1)
Departure time	T_D	T_D – 15 min.	T_D – 30 min.	T_D – 45 min.
Travel time	T_F	T_F + 15 min.	T_F + 30 min.	T_F + 45 min.
Arrival time	T_A	T_A	T_A	T_A
Toll	€ 6	€ 3	€ 1	€ 0

Notes: T_D stands for the departure time; T_A for the arrival time; and T_F for the travel time.

If the respondent chooses alternative C, we can infer that he is willing to pay €1 to save 15 minutes of travel time (C is preferred over D, implying a VOT of at least €4 per hour), but not more than €8 per hour (C is preferred over B indicating that the respondent is not willing to pay €2 to save 15 minutes). The calculation of the interval estimates based on the data will be explained later in Section 4.3. We have developed scenarios to estimate the other parameters in a similar way: these can be found in Appendix 4A. Note that the simplicity of the design in Table 4.1 comes at a price: we are not able to study the question of whether the VOT (and other valuations) depends on the size of the travel time gain.

The second part of the choice experiment consisted of 11 screens. Similar to the first part of choice experiment, respondents were also asked to allocate 10 trips over 4 alternatives here. The underlying design of this part of experiment was based on the conventional design used in the standard stated choice experiment[15] (see Louviere et al., 2000), which was different from the first experiment. The design has generated 44 choice sets, and was blocked into 4 sets of 11. Each respondent was assigned a block randomly, and the order of the 11 treatments in a block was randomized as well. The levels of attributes of the constructed alternatives are based on a fractional factorial design (orthogonal non-linear main effects design) using 4 levels for 13 of the attributes and 2 levels for two of the attributes. The attributes are based on the respondents' current behavior, with the aim to design alternatives as close to the reality of the individual respondent as possible (see Appendix 4B for an example). Each of the attributes has a limited number of values (levels), and these levels are combined in a systematic way, such that each attribute varies independently of the others. Each screen consists of 4 alternatives with separate attributes (alternative-specific attributes, see Table 4.2). Three alternatives are car-specific; the fourth alternative is always public transport (even in cases where the respondent indicated that there is no public transport alternative available; the choice sets concern hypothesized situations). The attribute levels are such that the first car alternative (A) is based on the preferred travel conditions of the respondent when paying a relatively high price. Expected arrival times are relatively close to the preferred arrival time, a large part of the trip is free flow, and uncertainty margins are small compared with alternatives B and C. The other road possibilities (alternative B and C) have lower road pricing fees but in return the travel conditions (in terms of arrival time, travel time, uncertainty and trip length in alternative C) are less attractive. The labels (in brackets in column 1) used in Table 4.2 to distinguish the three car trips were not shown to the respondents.

[15] The development was coordinated by the Technical University of Delft; for a detailed description of the experiment, we refer to Amelsfort (2004).

Table 4.2 Design of the second part of the stated choice experiment (11 screens)

Alternative	Attribute	Levels
A: car (pay)	Arrival time	4 (-10, -5, PAT, +5)[a]
	Travel time	4 (85% of trip length free flow, 90%, 95% and 100%)[b]
	Uncertainty	4 (uncertainty margin * 0.2, 0.4, 0.6 and 0.8)[c]
	Trip costs (fuel + charge)	4 (charge depends on distance, distance*0.08, 0.1, 0.12, and 0.14)
B: car (change departure time)	Arrival time	4 (-50, -30, -10, PAT,+10)[a]
	Travel time	4 (65% of trip length free flow, 70%, 75% and 80%) [b]
	Uncertainty	4 (uncertainty margin * 0.8, 1, 1.2 and 1.4)[c]
	Trip costs	4 (charge depends on distance, distance * 0.03, 0.04, 0.05, and 0.06)
C: car (change route)	Arrival time	4 (-30, -20, -10, PAT)[a]
	Travel time	4 (55% of trip length free flow, 60%, 65% and 70%) [b]
	Uncertainty	4 (uncertainty margin * 0.6, 0.8, 1, and 1.2)[c]
	Trip costs	4 (charge depends on distance, distance * 0, 0.01, 0.02, and 0.03)
	Trip length	2 (distance * 1.2, and 1.4)
D: public transport[d] (change mode)	Arrival time	4 (-30, -10, +10, +30 compared with PAT)[a]
	Travel time	2 (based on reported travel time with public transport if available (if not: 1.3*mean car travel time), no change, and reported travel time* 1.2)

Notes: [a]PAT = preferred arrival time.
[b]Travel time consists of a free-flow and a congested part.
[c] Uncertainty margin = difference between reported mean travel time and free-flow travel time.
[d]Trip cost is not one of the design attributes in public transport. That is, there is only one level of trip cost for public transport, and its actual level is based on each respondent’s trip length.

4.3 EXPERIMENT A: ANALYSIS OF INTERVAL ESTIMATES

4.3.1 How to derive interval estimates

The previous section described the interval values defined for experiment A, and the design of the first four screens. In order to calculate an average interval estimate for an individual, we need an expected value for each of the four intervals. It is not plausible to assume that the expected values are the exact middle points of their interval, and, besides, this is not defined for the fourth interval. We therefore hypothesize that there is an underlying statistical distribution that can be fitted to the actual aggregated trip allocation over the intervals, and approximate the expected interval values based on this presumed distribution. We have chosen to use the Gamma distribution. This distribution has been chosen for pragmatic reasons (in particular because only two parameters have to be estimated) and its non-negativity which is appropriate here. Note that the distribution is used only to determine the expected value for each interval, and it is not used as an approximation for the aggregate choice probabilities. In order to find the parameters of the best fitting Gamma distribution, we have applied the least squares method over the pooled choice probabilities (finding the minimum difference between the actual and the simulated distribution).

Table 4.3 Actual frequency of choices and the fitted Gamma distribution (parameters of the distribution in brackets)

VOT	**€ 0-4**	**€ 4-8**	**€ 8-12**	**> € 12**	**Total**
Actual probability	0.20	0.21	0.32	0.27	1.00
Gamma distribution (2.0; 0.2)	0.19	0.28	0.22	0.31	1.00
VSDE	**€ 0-2**	**€ 2-4**	**€ 4-6**	**> € 6**	**Total**
Actual probability	0.27	0.20	0.27	0.25	1.00
Gamma distribution (1.5; 0.32)	0.27	0.27	0.19	0.28	1.00
VSDL	**€ 0-8**	**€ 8-16**	**€ 16-24**	**> € 24**	**Total**
Actual probability	0.45	0.18	0.18	0.19	1.00
Gamma distribution (1.0; 0.07)	0.43	0.24	0.14	0.19	1.00
VUNC	**€ 0-3**	**€ 3-6**	**€ 6-9**	**> € 9**	**Total**
Actual probability	0.36	0.24	0.26	0.14	1.00
Gamma distribution (1.7; 0.31)	0.33	0.32	0.18	0.17	1.00

Table 4.3 shows the frequencies of choices over the four intervals and the Gamma distribution that could be fitted to these observations. With the parameters estimated, the expected value for each interval can be determined. Note that the probabilities from the Gamma distribution may deviate from the observed probabilities. It is therefore important to emphasize that the Gamma distribution was only used to determine the expected values within each interval for each choice. It appeared that the distributions were (slightly) different for income, so the mean interval value also depends on the income of the respondent. Table 4.4 presents the expected average values for the VOT, VSDE, VSDL and VUNC for each interval, for different income groups.

It is now possible to determine an individual's VOT as the weighted average of the intervals' expected values, where the individual's weights are determined by the number of trips allocated to each interval. For instance, when a respondent with an income of less than €28,500 allocates 5 trips to B and 5 trips to C, a VOT estimate of 7.8 results ((5*5.9+5*9.8)/10). The VSDE, VSDL and VUNC have been estimated in a similar way, only the interval values and attribute values were different (see Appendix 4a for the screens and interval values). Note that the interval boundaries calculated in this way refer to the average VOT, over discrete (non-marginal) travel time gains or losses. This is different from the conventional definition, which refers to the marginal rate of substitution between time and money.

Table 4.4 The expected values for the VOT, VSDE, VSDL and VUNC for each interval for the different income groups (€/hour).

Income (gross yearly)	**VOT**				**VSDE**				**VSDL**				**VUNC**			
	0-4	**4-8**	**8-12**	**>12**	**0-2**	**2-4**	**4-6**	**>6**	**0-8**	**8-16**	**16-24**	**>24**	**0-3**	**3-6**	**6-9**	**>9**
<€28,500	2.4	5.9	9.8	18.5	1.1	2.9	4.9	9.6	3.5	11.7	19.7	44.1	1.6	4.4	7.3	13.4
€28,500-45,000	2.4	5.9	9.8	18.1	1.1	2.9	4.9	9.5	3.4	11.6	19.6	40.2	1.6	4.4	7.3	13.1
€45,000-68,000	2.7	6.0	9.9	17.6	1.1	2.9	4.9	9.5	3.5	11.6	19.7	40.2	1.6	4.4	7.3	13.3
>€68,000	2.7	6.0	9.9	17.9	1.1	2.9	4.9	9.5	3.2	11.6	19.6	38.9	1.6	4.4	7.3	12.9

In principle, it is possible to analyze the choice data by discrete choice analysis. Nevertheless, the choice alternatives from this experiment are constructed in such a way that there is a serious multicollinearity problem among the attributes. Thus, there exists the problem either of model convergence or of having full sets of insignificant estimates in the resulting discrete choice models.

4.3.2 Results and statistical analysis

Table 4.5 shows the mean values for the various estimates. The mean value of time is about €10, which is somewhat higher than the current value that is used in policy documents, of about €8.3. The 95% confidence interval is between €9.6 and €10.1 which indicates that the difference is significant. The interval estimate of the VSDE is considerably lower than that of the VSDL. This confirms the expectation that people normally prefer early arrivals over late arrivals. The VUNC has a mean value of €5.4. As described in Chapter 2, it is difficult to compare this estimate with other international estimates, given the variety in the definitions of 'uncertainty' used. In this case, uncertainty is measured by the difference between the minimum and maximum travel times. It is evident that this ratio is different from when the units of uncertainty are based on the standard deviation or other measures of distribution (see also Chapter 2). The minimum and maximum values and the standard deviation derived from these interval estimates are shown in Table 4.5. Note that there is considerable variation among the respondents (especially for the VSDL).

The reported value of uncertainty in Table 4.5 is derived from the scenario presented in Appendix 4A. Uncertainty involves a minimum and maximum arrival time, and hence an (equal) chance of arriving earlier or later than the preferred arrival time. As a consequence, this value of uncertainty implicitly includes expected scheduling costs. It therefore makes sense to distinguish between two different concepts of the VUNC, with and without correction for expected scheduling costs, not only for conceptual reasons but also for comparability with the SCE discussed in Section 4.4. We can calculate the 'pure VUNC', labeled VUNC2 in Table 4.5, for each individual by reducing the present interval estimate (VUNC) by the estimated scheduling costs. For the calculation of these latter costs we use the individual's VSDE and VSDL. The probabilities of arriving earlier and later are 0.75 and 0.25, respectively. These probabilities are multiplied by 0.5 since the expected schedule delay, conditional on being early or late, is half the limit of the (uniform) distribution. So, the VSDE estimates are multiplied by 0.375 and the VSDL estimates by 0.125 in order to calculate the expected schedule delay costs. Both costs are then the expected schedule costs for being late and early, which are deducted from the VUNC for each individual. The average equals 1.82, considerably lower than the VUNC estimate, but still positive. The minimum value shows that for some respondents the pure VUNC may actually be negative, suggesting risk-loving.

Table 4.5 Descriptive statistics of interval estimates for the VOT, VSDE, VSDL, and VUNC (€/hour)

	N	Minimum	Maximum	Mean	Std. Deviation
VOT	1115	2.5	18.5	9.9	5.0
VSDE	1115	1.1	9.6	4.7	2.8
VSDL	1115	3.6	38.3	14.5	11.8
VUNC	1115	1.7	12.8	5.4	3.3
VUNC2	1115	-7.0	12.4	1.8	2.7

It is then interesting to search for explanatory variables. Since we have information on the socioeconomic characteristics of the respondent, it is possible to analyze their impact on the estimates. The literature indicates, for instance, that people with higher incomes tend to have a higher VOT. Table 4.6 shows the estimates for four different income categories. The results are, however, somewhat ambiguous. The highest income group does indeed have the highest VOT, but the high estimate for the lowest income group is more difficult to explain.

Table 4.6 The average values of the VOT, the VSDE, the VSDL, and the VUNC for the different income groups (€/hour)

	VOT	VSDE	VSDL	VUNC
< €28,500	9.9	4.6	18.6	5.8
€28,500-45,000	9.2	4.3	14.9	5.0
€45,000-68,000	9.8	4.7	13.6	5.3
> €68,000	10.5	5.0	12.6	5.2

The effect of income and other possibly important explanatory variables have been tested statistically. We conducted a regression analysis with the interval estimates as the dependent variable. Table 4.7 shows the results for the four regressions. Despite the very low overall fit of the models, the significance and sign of the coefficients give a tentative indication of the impact of the various variables. When we first look at income, the previous finding (in Table 4.6) for the VOT and the VSDE is confirmed: income has positively significant effect (at the 10% level) on the VOT and the VSDE. The effect of income is negatively significant on the respondents' VSDL. Income and education (also negatively significant) may be correlated here. A possible explanation for the negative impact of education and income on VSDL is that lower-educated people more often have jobs with less flexible working hours.

When we look at the results for VSDE, gender has a significant impact, with females having a higher VSDE than male respondents. Respondents who are higher-educated tend to have a lower VSDE. The explanation may be the same as for VSDL. The impact of travel-cost compensation by the employer (a higher VOT and VSDE for respondents who are fully-compensated) may be explained by the respondents' assumption that their employer will also pay for future road toll costs. The number of working hours during a week does not have a significant impact on any of the dependent variables. But the composition of the household may have some impact on the different concepts discussed here. Single-parent households with children are likely to have a higher VOT; families without children tend to have a higher VSDE, while single-person households seem to have a lower VUNC than other type of households.

Table 4.7 Regression results for the VOT, VSDE, VSDL and VUNC

	VOT		VSDE		VSDL		VUNC	
	B	t-stat.	B	t-stat.	B	t-stat.	B	t-stat.
Constant	8.122*	7.80	3.585*	6.11	21.73*	8.94	6.652*	9.66
Gender (dummy, female=1)	0.521	1.25	0.430*	1.84	0.233	0.24	0.352	1.28
Education	-0.164	-1.52	-0.138*	-2.26	-0.867*	-3.44	-0.156*	-2.18
Gross yearly income	0.120*	2.00	0.056*	1.67	-0.319*	-2.27	-0.056	-1.40
Cost comp1 (dummy)	-0.817	-1.59	-0.110	-0.38	2.206*	1.84	0.136	0.40
Cost comp2 (dummy)	-0.676*	-2.02	-0.373*	-1.97	0.054	0.07	-0.005	-0.03
Household composition1 (dummy)	-0.016	-0.04	0.176	0.70	-1.578	-1.51	-0.581*	-1.96
Household composition2 (dummy)	2.104*	2.22	0.566	1.06	-0.535	-0.24	-0.170	-0.27
Household composition3 (dummy)	0.536	1.54	0.351*	1.79	1.212	1.49	0.171	0.74
Working hours a week	0.036	0.28	0.087	1.20	0.179	0.60	-0.066	-0.78
R-square	.018		.018		.034		.015	

Notes: Significance is indicated by *, referring to significance at the 90% level.
Cost comp1: respondents receive no compensation from employer, cost comp2: respondents are partly compensated, cost comp3: respondents are fully compensated by employer (reference level); household composition1 = single; household 2 = single with child; household 3 = couple without children; household 4 = couple with child(ren) (reference level).

4.4 EXPERIMENT B: STATED CHOICE EXPERIMENT: THEORETICAL FRAMEWORK

4.4.1 Discrete choice models – Multinomial logit (MNL) model

Individual traveler's choice behavior is commonly analyzed using discrete choice models (e.g. Ben-Akiva and Lerman, 1987). Most models used in practice are based on the random utility theory (RUT), which assumes that individual's preference/taste can be described by a deterministic (systematic) part of utility, V_{ij}, and a stochastic component, ε_{ij}. The conventional random utility specification in the case of respondent *i* choosing among *J* alternatives is expressed in Eq. (4.1).

$$U\left(choice\ j\ for\ individual\ i\right) = U_{ij} = V_{ij} + \varepsilon_{ij}, \quad j = 1,\ldots, J. \tag{4.1}$$

The systematic component is assumed to be the part of utility contributed by attributes that can be observed by researchers, while the random component is the part of utility contributed by attributes unobserved by researchers. The observed part of systematic utility V_{ij} is a function of attributes in the alternative and the characteristics of the decision maker. A linear in-parameters function, which is specified by a vector of the decision maker's taste β, can be denoted as $\mathrm{V}_{ij} = \sum_{k=1}^{K} \beta_k X_{ijk}$ in the case of generic parameters (Ben-Akiva and Lerman, 1987). Utility maximization theory assumes that an individual chooses the alternative that yields the highest utility level.

The empirical specification of V_{ij} is crucial to modeling the individual's choice behavior because the utility function not only reflects an individual's decision making process given the socioeconomic environment, but also determines the predictive capability of the choice model.

In making the choice model operational, the random terms (unobserved by the analyst) also play a crucial role. Different assumptions on the joint distribution of random terms in the utility function result in different models. The most extensively used model in transportation studies is the Multinomial Logit (MNL) model, which assumes that the random terms are independently and identically distributed according to an extreme value type I distribution. Under these assumptions, the choice probability for respondent i to choose alternative j becomes:

$$\mathrm{Prob}_{ij} = \frac{\exp(V_{ij})}{\sum_{l=1}^{J} \exp(V_{ij})}. \tag{4.2}$$

This model can be solved using a maximum likelihood estimation method. The log-likelihood function is given as:

$$\log L = \sum_{i=1}^{I} \sum_{j=1}^{J} d_{ij} \log \mathrm{Prob}_{ij} . \tag{4.3}$$

In this present study, the dependent variable is the choice proportions allocated among four alternatives. Thus, d_{ij} is defined as the choice proportion distributed by respondent *i* to alternative *j*, and we have $\sum_j d_{ij} = 1$ under each choice profile (Greene, 2003). In our study, the respondent makes 11 choices in the experiment B.

4.4.2 Discrete choice models – Mixed logit (ML) model (ML)

The Mixed logit (ML) model generalizes the standard logit model by allowing the parameters to vary across individuals. Unlike the MNL model, which assumes the coefficients are homogeneous (fixed) among respondents, the ML accounts for the respondents' heterogeneity in taste by assuming the selected parameters follow some specific distribution(s) (McFadden and Train 2000). Furthermore, it also allows efficient estimation when there are repeated choices by the same individual (Revelt and Train, 1998). Observations drawn from the same individual are a common feature of SP data. The presence of multiple observations for each individual, implying the potential for correlated responses across observations, violates the independence-of-observations assumption in the classical MNL model. The ML models allow the model to be specified in such a way that the error components in different choice situations from a given individual are correlated, as occurs in our application.

To introduce the stochastic elements through β into the utility function, the taste of individual q is assumed as (Hensher and Greene, 2003):

$$\beta_{qk} = \beta_k + \delta_k' z_q + \eta_{qk}, \tag{4.4}$$

where η_{qk} is a random term whose distribution over individuals depends on underlying assumed distributional forms such as normal, lognormal, uniform, or triangular, and z_q is observed data. The paper of Hensher and Greene (2003) provides a comprehensive discussion on the state of the art of mixed logit models. The mixed logit estimations and computations in this and later chapters are carried out by the NLOGIT program.

4.5 EXPERIMENT B: ESTIMATION RESULTS OF STATED CHOICE EXPERIMENT DATA

The previous section described the basics of the maximum likelihood estimation of the utility parameters of the discrete choice models: the MNL and ML models. This section now discusses the application to our data. We study the tradeoffs between paying when traveling under preferred conditions versus paying less but facing less attractive travel conditions in terms of departure time, uncertainty, route length, travel time and mode. We estimated various specifications of this choice model, but can, of course, only present a selection of these. First, the basic model is outlined, including the resulting estimates of the VOT, VSDE, VSDL and VUNC. Next, we will include heterogeneity into the estimation of the models.

4.5.1 Multinomial logit (MNL) model

As a starting point, we analyze respondents' overall tradeoffs between mean travel time, uncertainty of travel time, and travel cost. This is similar to the 'mean-variance' modeling approach proposed by Jackson and Jucker (1981), where travelers were supposed to make a tradeoff between mean travel time and variance of travel time. This gives the estimates of how people evaluate travel time and uncertainty with respect to the monetary cost. The SCE included three different car alternatives and one public transport alternative. It makes sense to assume that all car alternatives have the same parameters, while those for public transport may be different. We have therefore defined two utility expressions: one for the car alternatives, and one for public transport. The utility expressions of car (V_C) and public transport modes (V_{PT}) are given by:

$$V_{CAR} = \beta_C \cdot C + \beta_\alpha \cdot E[T] + \beta_{UNC} \cdot UNC + EDT + VEDT + VVEDT \, ;$$

$$V_{PT} = ASC_{PT} + \beta_C \cdot C + \beta_\alpha \cdot E[T] + EDT + VEDT + VVEDT \, , \tag{4.5}$$

where *C* is the travel cost (in our experiment consisting of both fuel and toll costs for the car, and ticket price for public transport), *E[T]* is the mean travel time; and *UNC* is the amount of uncertain travel time[16]. ASC_{PT} is the alternative-specific constant of public transport, added to capture the effect of respondents' difference in preference for car or public transport. Our experiment also involves different departure time conditions, implied by different mode and route alternatives. In this model, we allow for nonlinear schedule delay early effects, and we specify a set of dummy variables, *EDT*, *VEDT*, and *VVEDT*, to explain the utility difference incurred by different chosen departure time slots. *EDT* denotes the dummy for 'early departure' and captures the disutility when the respondent's departure time is 30 to 60 minutes earlier than his/her

[16] The mean travel time is defined as the mean value of minimum and maximum total travel time in the choice experiment, while uncertainty is the difference between maximum and minimum total travel time.

preferred departure time (PDT); *VEDT* is the dummy of 'very early departure', accounting for the disutility when the respondent's departure time is 60-90 minutes earlier than his/her *PDT*; and *VVEDT* gives the 'very very early departure' dummy when the departure time is more than 90 minutes earlier than the *PDT*. Here, we use effects coding for the early departure variables of EDT, VEDT, and VVEDT. This implies that, instead of coding the base level[17] 0 across the newly created variables, the base level is -1 across each of these new variables. The purpose of using effects coding is to avoid the perfect correlation between the effect of the base level and the effect of intercept in the model (i.e. alternative-specific constant of public transport 'ASC_{PT}' in this case).

Next, we estimate a more complete model incorporating the scheduling variables based on Eq.(2.5). This model assumes that the individual accounts for the following attributes in their decision making: travel cost, *C*; mean travel time *E[T]*; expected schedule delay early, *E[SDE]*; expected schedule delay late, *E[SDL]*; probability of arriving later than the preferred arrival time, P_L[18]; and amount of uncertain travel time (*UNC)*. The generic indirect utility functions of car (V_{CAR}) and public transport modes (V_{PT}) are given as follows:

$$\begin{aligned} V_{CAR} &= \beta_C \cdot C + \beta_\alpha \cdot E[T] + \beta_\beta \cdot E[SDE] + \beta_\gamma \cdot E[SDL] + \beta_\theta \cdot P_L + \beta_{UNC} \cdot UNC \\ &\quad + EDT + VEDT + VVEDT; \\ V_{PT} &= ASC_{PT} + \beta_C \cdot C + \beta_\alpha \cdot E[T] + \beta_\beta \cdot E[SDE] + \beta_\gamma \cdot E[SDL] + \beta_\theta \cdot P_L \\ &\quad + EDT + VEDT + VVEDT. \end{aligned} \tag{4.6}$$

Uncertainty of travel time is also included as an independent variable in this model. However, it is likely to be of limited relevance, since most of the uncertainty effects are captured by *E(SDE)*, *E(SDL)*, and P_L (recall the discussion about VUNC and VUNC2 in Section 4.3).

The first two columns of Table 4.8 show the MNL estimates of the mean-variance modeling and the trip scheduling modeling approach. The unit of all time-related attributes is in minutes, and travel cost is in euros. A general finding obtained from these two models is the negative sign for

[17] The base level of early departure variables is set as: the departure time is less than 30 minutes earlier than the PDT.
[18] The computation of E[SDE], E[SDL], and P_L is given in Appendix 4C.

the alternative-specific public transport coefficient. This suggests that, in general, respondents prefer the car over the public transport alternative, because the ASC represents individuals' taste in choosing that alternative. A possible explanation for this phenomenon is that the samples consist of car users, who implicitly have a preference for the car. Shifting departure time to earlier time periods is not considered attractive in all models.

Compared to Model 1, Model 2 has improved the log likelihood substantially (from -15,549.72 to -15,420.39). Using the likelihood ratio (LR) test at any generally acceptable level of confidence, we can confidently reject the null hypothesis of no significant difference between Model 1 and Model 2. This indicates that the use of trip scheduling modeling approach has a better goodness-of-fit of the model than the use of mean-variance modeling approach. Note that that the effect of uncertainty is only significant in Model 1. In Model 2, with the work scheduling consideration taken into account, uncertainty is no longer significant. This result suggests that the value of uncertainty may be explained by its effect on expected schedule delay costs. Small et al. (1999) obtained a similar result for the estimate of the standard deviation of travel time in the scheduling specification utility function. These authors argue that, when the scheduling costs are fully specified in a model, it is probably unnecessary to add an additional cost for unreliability (uncertainty) of travel. Our results confirm this.

Based on the specification in equation (4.6), we extend our analysis and investigate the mode-specific effects by interacting the travel time and scheduling variables with a public transport dummy. This leads to the following general specification:

$$\begin{aligned} V = {} & PT \cdot ASC + \beta_C \cdot C + \beta_\alpha \cdot E[T] + \beta_\beta \cdot E[SDE] + \beta_\gamma \cdot E[SDL] + \beta_\theta \cdot P_L + \beta_{UNC} \cdot UNC \\ & + \beta_{PT\alpha} \cdot PT \cdot E[T] + \beta_{PT\beta} \cdot PT \cdot E[SDE] + \beta_{PT\gamma} \cdot PT \cdot E[SDL] \\ & + EDT + VEDT + VVEDT, \end{aligned} \tag{4.7}$$

where $PT = 1$ if the alternative is public transport, and zero otherwise.

The aim is to analyze whether respondents evaluate the attributes of public transport and road transport differently. By checking the significance of the coefficients of these interaction terms, we can examine whether the valuation of public transport significantly differs from road transport. The results of Model 3 (Table 4.8) do indeed show that there is a difference in the disutility attributed to travel time and schedule delay. Surprisingly, travel time becomes more important and scheduling deviations less important for public transport relative to the car.

Finally, we considered the nonlinear effects of scheduling variables, such as *E[SDE]* and *E[SDL]*, by including the quadratic terms of these variables in our indirect utility function (Models 4 and 5). The coefficient of this quadratic term of SDE is negative and significant, indicating a nonlinear effect. It indicates that people's aversion to arriving early is increasing more than proportionally as their schedule delay early time increases. Based on the findings of Model 4, the public transport interaction terms are re-inserted in Model 5. Because there is almost no explanatory power of the variables *E[SDE]* and $E[(SDL)^2]$ in Model 4, we omit these two variables in Model 5.

The resulting parameter values (VOT, VSDE, VSDL and VUNC) from these models are summarized in Table 4.9. The generic VOT values around 8.9 €/hour seem reasonable and are in-between the values used in Dutch policy documents and the (mean) interval estimates presented in Section 4.3. Similar results are found for the VSDL. The value of uncertainty is somewhat lower than the interval estimate (only significant when scheduling costs are excluded).

The results for Model 2 indicate a considerably higher estimate for VSDE than in the interval estimate. This value (12.01 €/hour) is even higher than the VOT, which is in fact rather implausible. People would then prefer extra travel time over early arrival at work. Although this may sometimes be true, it is not something one would expect to be true on average, over all respondents. The high VSDE value may be explained by the nonlinear effect of the SDE variable.

We have seen that inclusion of the quadratic terms of SDE (Model 4) leads to a coefficient of $E[(SDE)^2]$ that is negative and significant. Because expected SDE appears as a quadratic term in the utility function, the marginal cost of SDE rises with SDE. Consequently, the VSDE is within a reasonable range when the expected schedule delay early time is within 20 minutes. This finding is plausible, as similar results were also obtained in previous studies (Hendrickson and Plank, 1984; Small et al. 1999). It is also in line with the interval estimate.

Table 4.8 Estimation results of the basic models for SCE data

Explanatory variables	Model 1	Model 2	Model 3	Model 4	Model 5
ASC of public transport alternative A_PT	-1.068*** (-22.04)	-0.772*** (-14.85)	-0.784*** (-7.78)	-0.7783*** (-14.944)	-0.7554*** (-8.220)
Travel cost C	-0.095*** (-16.69)	-0.094*** (-16.54)	-0.095*** (-16.64)	-0.0936*** (-16.466)	-0.0946*** (-16.521)
Mean travel time E[T] E[T]	-0.014*** (-11.49)	-0.014*** (-9.58)	-0.013*** (-8.87)	-0.0139*** (-9.588)	-0.0132*** (-8.839)
E[T]*Public transport dummy E[T]*PT			-0.002** (-2.20)		-0.0020** (-2.131)
Expected schedule delay early E[SDE]		-0.019*** (-8.58)	-0.020*** (-8.70)	-0.0080 (-1.189)	
Expected schedule delay early squared $E[(SDE)^2]$				-0.0002* (-1.653)	-0.0004*** (-8.650)
E[SDE]*Public transport dummy E[SDE]*PT			0.009** (2.36)		
Expected schedule delay early squared*public transport dummy $E[(SDE)^2]$*PT					0.0003** (2.269)
Expected schedule delay late E[SDL]		-0.023*** (-9.42)	-0.026*** (-9.24)	-0.0270*** (-5.595)	-0.0262*** (-9.623)
Expected schedule delay late squared $E[(SDL)^2]$				0.0001 (0.967)	
E[SDL]*Public transport dummy E[SDL]*PT			0.010** (2.50)		0.0093** (2.542)
Probability of late arrival (later than PAT) P_L		-0.092* (-1.89)	-0.072 (-1.46)	0.0071 (0.099)	0.0564 (1.301)
Uncertainty UNC	-0.007*** (-5.25)	0.002 (1.30)	0.002 (1.00)	0.0016 (1.130)	0.0018 (1.153)
Dummy for departing 30-59 min earlier than preferred departure time (PDT) EDT	0.346*** (12.06)	0.294*** (8.94)	0.293*** (8.72)	0.2860*** (8.574)	0.2817*** (8.252)
Dummy for departing 60-89 min earlier than PDT VEDT	-0.209*** (-6.38)	-0.139*** (-3.74)	-0.146*** (-3.85)	-0.1335*** (-3.573)	-0.1446*** (-3.808)
Dummy for departing more than 90 min PDT VVEDT	-0.533*** (-7.35)	-0.521*** (-5.96)	-0.516*** (-5.81)	-0.5237*** (-5.985)	-0.5318 (-6.028)
Log likelihood Pseudo-R-sqrd	-15549.72 0.085	-15420.39 0.093	-15413.53 0.093	-15418.60 0.093	-15413.53 0.093

Notes: t-statistics are shown in parenthesis. Significance is indicated by ***, **, and * referring to significance at the 99%, 95%, and 90% level, respectively.

The estimated coefficients of the interaction terms in Model 3 and Model 5 indicate that the valuations of travel time and scheduling attributes in public transport are significantly different from those in car transport. Figures for the VOT, derived from Models 3 and 5, are significantly higher for public transport than for road transport. Jiang and Morikawa (2004) theoretically analyzed the variation of VOT, and they found that the value of travel time savings is higher for a slower mode if the marginal disutility of time losses increases with travel time. This is the case in our experiment. Public transport is designed as a slower mode and marginal utility is likely to decrease when travel time rises. Another possible explanation is that public transport is generally less comfortable, especially in the peak; people may then be willing to pay relatively more to reduce public transport travel time than time spent in a car. In this context, it is important to emphasize that we are dealing here with valuations by car users. For the VSDE and VSDL, car transport has (significantly) higher estimates than public transport. A possible explanation is that public transport usage may be perceived as a relatively easy excuse to arrive early or late at work. Private car users may therefore be willing to pay more to save early or late arrival time. It may also be that people, because they can distribute 10 trips over 4 alternatives, have in mind trips with relatively small scheduling concerns for public transport.

Table 4.9 Monetary values of time and other time attributes of Models 1-5(€/hour)

	Model 1	Model 2	Model 3	Model 4	Model 5
VOT generic	8.87	8.88	-	8.94	-
VOT for car	-	-	8.35*	-	8.35*
VOT for public transport	-	-	9.66*	-	9.63*
VSDE generic	-	12.01	-		-
At SDE=10 min				3.17	
At SDE=20 min				6.34	
At SDE=30 min				9.51	
VSDE for car	-		12.30*	-	
At SDE=10 min					5.52*
At SDE=20 min					11.04*
At SDE=30 min					16.56*
VSDE for public transport	-		6.43*	-	
At SDE=10 min					2.29*
At SDE=20 min					4.59*
At SDE=30 min					6.88*
VSDL generic	-	14.58	-	17.29	-
VSDL for car	-	-	16.07*	-	16.60*
VSDL for public transport	-	-	9.69*	-	10.73*
VUNC generic	4.24	-	-	-	-

Notes: All monetary values given in this table are significant at the 90% level; * indicates that the difference between car and public transport is significant at the 90% level.

4.5.2 Observed heterogeneity: multinomial logit model with a set of covariates

This section elaborates the analysis by allowing the travel time and scheduling-related attributes to interact with behavioral indicators, such as restriction of work starting time and restriction of home departure time, and with some socioeconomics indicators, such as gender, income, education, and travel cost compensation. The sample characteristics of these variables can be found in Appendix 4F. The covariate effects investigated in this subsection are all based on the utility specification in Model 2 (Eq.(4.6)).

Behavioral indicators: effects of departure and arrival time restrictions

Intuitively, an individual's flexibility of arrival time at work and departure time from home will have some impact on the VOT and scheduling costs. Numerous empirical studies have confirmed that work starting time flexibility has a significant impact on the schedule delay estimates (e.g. Small, 1982; Small et al., 1999). Most studies have focused on arrival time restrictions; fewer studies have explicitly addressed the impact of departure time flexibility on schedule delay costs. Our data contains information on both issues. We have therefore specified the interaction terms for these restriction dummies with time and scheduling attributes, and analyze the significance of both effects.

Table 4.10 shows the VOT, VSDE, and VSDL (the underlying MNL models can be found in Appendix 4D). It appears that restrictions do matter, as may be expected. Respondents with restricted starting times at work have a significantly higher VSDE and VSDL, and they also incur a penalty for arriving later than the restricted time (included by means of a dummy variable). These respondents are willing to pay about €6.4 to arrive before the restrictive departure constraints. Commuters tend to have a higher VSDE when it is impossible to change their departure time to an earlier time slot. This is an indication that what is measured as a schedule delay early cost should preferably be interpreted in terms of scheduling problems with the preceding activity. This issue will be addressed at greater length in the next chapter. Commuters who cannot change departure time to a later moment have a higher VOT.

Table 4.10 Monetary values implied by Model 10 (shown in Appendix 4D) (€/hour)

	VOT	VSDE	VSDL	Penalty (later or earlier than restriction)
No restriction [a]	8.47	9.80	11.25	-
Cannot arrive at work later than restriction	7.57	12.61*	15.34**	6.39***
Cannot depart from home before restriction	10.29	18.29***	11.22	1.49
Cannot depart from home after restriction	13.25***	7.87	10.57	3.13**

Notes: ***, **, and *, indicate that the difference between one particular group and reference group are significant at the 99%, 95%, and 90% levels, respectively.
[a] No restriction on departure and arrival time is taken as the reference group for comparison.

Travel environment and socioeconomic indicators

There is ample evidence that the VOT, VSDE, and VSDL vary with travel environment and socioeconomic variables such as trip length, income, and gender (see, e.g., Small et al., 1999; Lam and Small, 2001). Here, we investigate the effects of trip length, income, gender, education, and travel cost compensation by the employer on our estimates of interest. The estimation results can be found in Appendix 4E, while the summarized monetary values are given in Table 4.11.

Most of the findings are in line with the literature, such as a positive trip-length effect on the VOTs (Gunn, 2001); positive income effects on the VOTs (Small et al., 1999) and positive gender (female) effects on schedule delay cost (Lam and Small, 2001). In particular, we also find that scheduling costs are lower for respondents with a higher income and a higher educational level (similar to what we found with the interval estimates). The reason for this may be that higher-educated people have higher-status jobs, which generally have less restricted working times. The impact of income on the VOT is different from the interval estimates where we did not find a positive relation. The compensation of travel costs by the employer seems most relevant for the VOT. Compensated commuters tend to have a higher VOT than uncompensated drivers. The results of the interval estimates were less strong on this issue; there was only a difference between fully- and partially-compensated respondents.

Table 4.11 Monetary values implied by Model 11 to 15 (€/hour)

	VOT	VSDE	VSDL
Trip length 30 km or less [a]	6.66	14.85	19.57
Trip length 30-60 km	6.57	9.51***	10.91***
Trip length 60 km or more	11.09***	11.11*	9.06***
Household yearly income 28,500 or less [a]	5.29	14.27	18.47
Household yearly income 28,500-45,000	6.49	11.22	16.48
Household yearly income 45,000-68,000	12.72***	9.72**	10.26***
Household yearly income 68,000 or more	10.48***	12.35	11.74***
Male [a]	8.51	10.35	15.02
Female	10.72**	17.31***	14.37
Lower education (HAVO or less) [a]	8.91	11.20	16.94
Higher education (HBO or above)	8.95	13.24	12.05**
No travel cost compensation [a]	- [b]	11.00	15.74
Partial travel cost compensation	9.74***	8.54	10.88*
Full travel cost compensation	12.50***	13.03	13.06

Notes: ***, **, and *, indicate that the difference between one particular group and reference group are significant at the 99%, 95%, and 90% level, respectively.
[a] This is taken as the reference group for comparison.
[b] The travel time coefficient is not significant even at the 90% confidence level.

4.5.3 Mixed logit models

The mixed logit (ML) model estimated in this section was described in Section 4.4.2. ML models are estimated with simulated log-likelihood so that simulations are needed in the estimation procedure. For each random parameter, we used 250 Halton drawings to calculate the simulated likelihood function. The model is assumed to be converged if the first derivative for all parameters is smaller than 10^{-4}.

The ML models presented in this section are all based on the utility specification as in Model 5 in the MNL case, and with four random parameters – expected travel time, expected schedule delay early squared, expected schedule delay late, and uncertainty. Here, we only presented the results of ML models incorporating the panel structure of the data set, which is the feature of the SCE where the repeated choices were observed from the same individual. The first ML model (ML1) we estimated is based on the assumption that all four random variables are normally-distributed. The second ML model (ML2) assumes of four triangular-distributed variables. Since the results of the normal and the triangular distributions include a large proportion of behaviorally implausible signs for the travel time attribute (i.e. a positive coefficient for the travel time variable), it may be appealing if we impose a constraint on the travel time variable. Thus, different from ML2, the third model (ML3) assumes that the parameter of expected travel time is constraint triangular, that is, it restricts the coefficient to be non-positive, and the other variables remain non-constraint triangular-distributed. We also estimated the model (ML4) where the expected travel time variable is lognormally distributed and the others are normally distributed. Table 4.12 summarizes the results of these four ML models. The implied mean VOT, VSDE, VSDL, and VUNC for car mode are also provided in Table 4.12.

Compared to the MNL model (Model 2 in Table 4.8), the ML models (given any model in Table 4.12) improved the pseudo R-squared substantially. The result of LR test shows that there is a major improvement for model fit of the ML models at any accepted level of statistical significance. This indicates the importance of including unobserved heterogeneity and incorporating the panel data structure from a data fit perspective. A comparison of the implied monetary values across the ML models (see the bottom rows of Table 4.12) shows that the variations of these monetary values, particularly the VOT and the VUNC, are dramatic. Though

ML1 and ML2 give more plausible values for VSDE and VSDL, when compared with the MNL in Model 5, the VOT values obtained from these two models are rather doubtful. In the cases of imposing constraints on the travel time parameter, i.e, ML3 and ML4, we found that the VOT values become more reasonable, whereas the VSDL and the VUNC values are biased downwardly in these two models.

We also tried different model specifications in the context of ML models: for example, specifying more, fewer, or different random parameters, different utility specifications, or incorporating some other observed heterogeneities (i.e. individual and trip characteristics) into the random parameters, and other variations, but we were not able to find any systematic pattern of the variations in the estimates. It is clear that our estimates are highly sensitive to the model specification in the context of the ML framework. This is a somewhat discomforting conclusion, and we hypothesize that the models tested here may not adequately reflect the respondents' true preference structures. We will test this hypothesis by proposing a different model in the next chapter.

Table 4.12 Estimation results of the ML models

Explanatory variables	ML 1 Normal distr. for all random parameters		ML 2 Triangular distr. for all random parameters		ML 3 Constraint triangular distr. for E[T]_Car and triangular distr. for the other parameters		ML 4 Lognormal distr. for E[T]_Car & normal distr. for the other parameters	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
Random parameter mean effects								
E[T]_Car	-0.0036	-2.53	-0.0013	-0.89	-0.0498	-45.05	-3.8487	-41.70
$E[(SDE)^2]$_Car	-0.0007	-12.20	-0.0007	-12.20	-0.0012	-20.47	-0.0010	-12.21
E[SDL]_Car	-0.0300	-13.30	-0.0295	-13.56	-0.0241	-8.87	-0.0269	-6.48
Uncertainty	-0.0094	-3.24	-0.0099	-6.66	0.0149	9.92	0.0074	2.08
Random parameter standard deviation								
E[T]_Car	0.0600	77.72	0.1463	76.24	0.0498	45.05	0.8641	23.31
$E[(SDE)^2]$_Car	0.0009	24.60	0.0021	21.69	0.0021	17.23	0.0009	12.34
E[SDL]_Car	0.0410	34.18	0.0996	49.36	0.1024	23.49	0.0404	13.28
Uncertainty	0.0166	24.19	0.0579	34.04	0.1035	31.42	0.0448	16.12
Non-random parameters								
E[T]_PT	-0.0386	-24.84	-0.0183	-25.88	-0.0570	-35.60	-0.0500	-28.07
$E[(SDE)^2]$_PT	-0.0004	-3.85	-0.0004	-3.90	-0.0007	-6.63	-0.0006	-3.11
E[SDL]_PT	-0.0185	-5.97	-0.0934	-5.96	-0.0027	-0.90	-0.0100	-1.87
P_L	0.0470	1.16	0.0454	1.12	0.1009	2.47	0.0827	1.29
ASC_PT	-0.4388	-4.28	-0.3464	-3.39	-0.3817	-3.87	-0.1679	-1.47
Travel cost	-0.0976	-18.63	-0.0957	-18.29	-0.1600	-27.19	-0.1416	-34.16
EDT	0.2237	7.09	0.2198	7.02	0.0675	2.05	0.1549	4.21
VEDT	-0.2017	-8.43	-0.1954	-8.03	0.0328	0.88	-0.0613	-1.37
VVEDT	-0.4094	-5.49	-0.4109	-5.58	0.2278	2.75	-0.1020	-1.05
Log-likelihood	-13754.31		-13780.32		-14407.18		-14222.52	
Pseudo-R-sqrd	0.1907		0.1892		0.1523		0.1631	
Mean monetary values (€/hour)								
VOT_CAR	**2.22**		0.80 (insignificant)		**18.67**		**9.03**	
VSDE_CAR (when SDE=10 min)	**4.34**		**4.39**		**4.42**		**4.14**	
VSDL_CAR	**18.47**		**18.49**		**9.04**		**11.37**	
VUNC_CAR	**5.75**		**6.21**		**-5.60**		**-3.13**	

Note: Bold print is used for significant values at the 90% level.

4.6 CONCLUDING REMARKS

Policy makers are usually very interested in up-to-date VOT estimates of road users. This concept is an important input in the cost-benefit analysis of new roads. Recently, with the increasing levels of congestion, other concepts have also gained in importance in the policy arena. Reliability has recently been proposed as an important indicator for the performance of the Dutch road infrastructure. In addition, the estimated choice models can be integrated in models which predict car drivers' responses to changing travel times or to the imposition of road-user charging.

The aim of this chapter was to find tradeoffs between paying and traveling under attractive conditions (in terms of arrival time, travel time, etc.) versus paying less (or nothing) but facing less attractive travel conditions in terms of departure time, route length, and mode. We have estimated choice models to estimate values for the above-mentioned concepts and assessed the impact of individual characteristics on these values.

Two different types of choice experiments were carried out. The first, relatively simple, choice experiment was developed to find individual (interval) estimates for each type of parameter value. In line with other empirical results, we found that the VSDL has the highest value (a mean value of 14 €/hour), followed by the VOT (about 10 €/hour). However, this latter value is slightly higher than the value used in 2004 for commuting traffic (8.3 €/hour) by Dutch policy makers. Commuters seem to attach less value to arriving early and uncertain travel times. The importance of socioeconomic characteristics is rather modest. Income, often cited as important for the VOT, only explains some of the variation in the VSDL estimate. Whether employers compensate for travel costs or not influences some of the parameter values. The VOT and the VSDE tend to be higher for fully-compensated respondents. Those respondents probably think that the road tolls will also be compensated by their employer.

The second experiment consisted of 11 choice sets, and again respondents were asked to allocate 10 trips over four alternatives. It was a labeled experiment in which the alternatives consisted of different attributes (15 in total), which were based on the current behavior of each individual. We estimated various choice models by using the choice proportions setup. The respondents preferred the car alternative over the public transport alternative. When we include scheduling costs into the estimations, “pure uncertainty” becomes insignificant. This has also been found by others, and it suggests that it may often be unnecessary to add an additional cost for unreliability (or uncertainty) of travel when scheduling costs are fully specified. The nonlinear effects of scheduling variables have also been addressed in our model estimations. The analysis indicates that people’s aversion to arriving early increases nonlinearly as their schedule delay early time increases.

The resulting parameter values for VOT and VSDL seem rather plausible. The size is comparable to the interval estimates, with the VOT being somewhat lower (comparable with the previously mentioned 2004 value used in policy documents). The generic VSDE estimates for car and public transport were rather high. This may be explained by the nonlinear effect of the SDE variable. VSDE decreases (to a more reasonable value) when the expected schedule delay early time is within 20 minutes. Note that the VOT, VSDE, and VSDL of public transport are significantly different from those in car transport. The VOT of car users is lower, whereas their VSDE and VSDL are significantly higher.

We have also included personal variables in the utility functions to study observed heterogeneity. It was found that the presence of departure or arrival time restrictions is important for the parameter values. People with restricted starting times at work have a higher VSDE and VSDL, and they also incur a penalty for arriving later than the restricted time. For the restrictions of individuals’ commuting departure time, the effects are different between early and late departure constraints. Moreover, trip length seems to have a significant impact on the VOT, VSDE, and VSDL. Respondents making longer commuting trips generally attach a lower value to arriving earlier or later than the preferred time. The VSDL is lower for people with a higher income, which is similar to the findings for the interval estimates. Travel cost compensation only seems to have an impact on the VOT, with fully-compensated commuters generally having a higher VOT. This is similar to what we found with the first experiment. These respondents may assume that future road charges will also be compensated by their employers, which can explain their willingness to pay for travel time savings.

The experience of estimating the ML models strongly indicates the existence of considerable heterogeneity among individuals; nevertheless, the less robust ML results may imply that the current scheduling model we used could not perfectly explain the travelers' choice behavior in the current data set. We will discuss this issue and propose another modeling framework in the next chapter.

Appendix 4A: Scenarios to obtain VSDL, VSDE and VUNC interval estimates

The literature suggests that the **VSDE** is about half of the VOT. Therefore, we defined the following 4 intervals:

1. € 0 – 2
2. € 2 – 4
3. € 4 – 6
4. > € 6

	A (group 4)	B (group 3)	C (group 2)	D (group 1)
Departure time	T_D	T_D – 15 min.	T_D – 30 min.	T_D – 45 min.
Travel time	T_f	T_f	T_f	T_f
Arrival time	T_A	T_A – 15 min.	T_A – 30 min.	T_A – 45 min.
Toll	€ 3	€ 1.50	€ 0.50	€ 0

According to the literature, the **VSDL** is about twice the VOT. Therefore, we defined the following 4 intervals:

1. € 0 – 8
2. € 8 – 16
3. € 16 – 24
4. > € 24

	A (group 4)	B (group 3)	C (group 2)	D (group 1)
Departure time	T_D	T_D + 10 min.	T_D + 20 min.	T_D + 30 min.
Travel time	T_f	T_f	T_f	T_f
Arrival time	T_A	T_A + 10 min.	T_A + 20 min.	T_A + 30 min.
Toll	€ 8	€ 4	€ 1.33	€ 0

We have defined, rather arbitrarily, the following intervals for the **VUNC**:

1. € 0 – 3
2. € 3 – 6
3. € 6 – 9
4. > € 9

	A (group 4)	B (group 3)	C (group 2)	D (group 1)
Departure time	T_D – 30 min.	T_D – 30 min.	T_D – 30 min.	T_D – 30 min.
Min. travel time	T_f + 30 min.	T_f + 5 min.	T_f + 0 min.	T_f
Max. travel time	T_f + 30 min.	T_f + 35 min.	T_f + 40 min.	T_f + 55 min.
Min. arrival time	T_A	T_A – 15 min.	T_A – 30 min.	T_A – 45 min.
Max. arrival time	T_A	T_A + 5 min.	T_A + 10 min.	T_A + 15 min.
Tol	€ 6	€ 3	€ 1	€ 0

Appendix 4B: Example of one screen (with 4 alternatives) of the second part of the SC-experiment, as presented to the respondent (levels are indicative)

Alternative A	Alternative B	Alternative C	Alternative D
Mode of transport: car	Mode of transport: car	Mode of transport: car	Mode of transport: public transport
Trip length : 35 km	Trip length: 35 km	Trip length: 49 km	Trip length: 35 km
Travel costs: € 8.10 of which:	Travel costs: € 4.60 of which:	Travel costs: € 6.20 of which:	Price of a ticket: € 3.18
– fuel: €3.20	– fuel: €3.20	– fuel: €4.20	
– charge: €4.90	– charge: €1.40	– charge: €2.00	
Departure time: 08.10	Departure time: 08.25	Departure time: 08.00	Departure time: 07.25
Total travel time between 40 and 50 minutes of which:	Total travel time between 50 and 60 minutes of which:	Total travel time between 55 and 65 minutes of which:	Total travel time: 72 minutes
– free flow: 25 min.	– free flow: 25 min.	– free flow: 40 min.	
– minimum time in congestion: 15 min.	– minimum time in congestion: 25 min.	– minimum time in congestion: 15 min.	
– maximum time in congestion: 25 min.	– maximum time in congestion: 35 min.	– maximum time in congestion: 25 min.	
Arrival time is hence between: 8.50 and 9.00	Arrival time is hence between: 9.15 and 9.25	Arrival time is hence between: 8.55 and 9.05	Arrival time: 08.37
Number of trips …..	Number of trips ….	Number of trips …	Number of trips …

Appendix 4C: Computation of E[SDE] and E[SDL]

SDE is defined to be positive for early arrivals, and zero otherwise; while SDL is positive for late arrivals, and zero otherwise. P_L represents the probability of arriving later than the preferred arrival time.

$$SDE(AT) = \max\{PAT - AT, 0\}$$

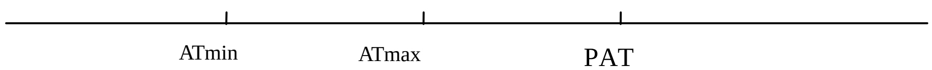
$$SDL(AT) = \max\{AT - PAT, 0\}$$

$$P_L = \Pr ob(AT > PAT),$$

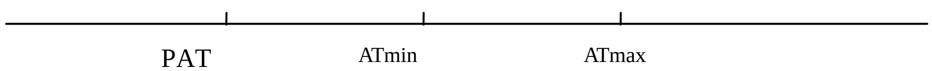
where AT denotes the arrival time, and PAT is the preferred arrival time.

To compute the E[SDE], E[SDL], and P_L, we can distinguish the following three cases:

Case 1:



Case2:



Case3:

where ATmin is the earliest arrival time, and ATmax is the latest arrival time

Case 1: $AT\max \le PAT$

$$E[SDE] = PAT - \frac{1}{2}\left(AT\min + AT\max\right)$$

$$E[SDL] = 0$$

$$P_L = 0.$$

Case 2: $AT\min \ge PAT$

$$E[SDE] = 0$$

$$E[SDL] = \frac{1}{2}\left(AT\min + AT\max\right) - PAT$$

$$P_L = 1.$$

Case 3: $AT\min < PAT \ \& AT\max > PAT$

$$E[SDE] = \frac{1}{2}\left(PAT - AT\min\right) * \left(1 - P_L\right)$$

$$E[SDL] = \frac{1}{2}\left(AT\max - PAT\right) * P_L$$

$$P_L = \frac{AT\max - PAT}{AT\max - AT\min}\ .$$

Appendix 4D: Estimation results of scheduling restriction effects based on Model 2

Explanatory variables	Model 6	Model 7	Model 8	Model 9	Model 10
ASC of public transport alternative	-0.7737***	-0.7747***	-0.7682***	-0.7726***	-0.7770***
	(-14.871)	(-14.872)	(-14.776)	(-14.858)	(-14.911)
Travel cost C	-0.0937***	-0.0938***	-0.0933***	-0.0942***	-0.0939***
	(-16.463)	(-16.457)	(-16.416)	(-16.581)	(-16.442)
E[T]	-0.0136***	-0.0143***	-0.0132***	-0.0132***	-0.0133***
	(-9.278)	(-8.952)	(-8.945)	(-8.988)	(-8.200)
E[T]*arriving later than work restriction		0.0009			0.0014
		(0.671)			(1.106)
E[T]*departing earlier than home restriction			-0.0051***		-0.0028
			(-2.698)		(-1.285)
E[T]*departing later than home restriction				-0.0058***	-0.0075***
				(-3.120)	(-3.762)
E[SDE]	-0.0186***	-0.0166***	-0.0170***	-0.0192***	-0.0153***
	(-8.491)	(-6.376)	(-7.604)	(-8.549)	(-5.773)
E[SDE]*arriving later than work restriction		-0.0040			-0.0044*
		(-1.572)			(-1.728)
E[SDE]*departing earlier than home restriction			-0.0145***		-0.0133***
			(-3.704)		(-3.315)
E[SDE]*departing later than home restriction				0.0025	0.0030
				(0.669)	(0.804)
E[SDL]	-0.0211***	-0.0176***	-0.0233***	-0.0223***	-0.0176***
	(-8.607)	(-6.089)	(-9.289)	(-8.955)	(-5.913)
E[SDL]*arriving later than work restriction		-0.0103***			-0.0064**
		(-3.285)			(-1.980)
E[SDL]*departing earlier than home restriction			0.0032		0.0001
			(0.702)		(0.011)
E[SDL]*departing later than home restriction				-0.0037	0.0011
				(-0.778)	(0.218)
Probability of late arrival (later than PAT)	-0.0940*	-0.0861*	-0.0950*	-0.0917*	-0.0936*
	(-1.932)	(-1.770)	(-1.953)	(-1.886)	(-1.920)
Uncertainty	0.0017	0.0017	0.0019	0.0018	0.0016
	(1.220)	(1.226)	(1.354)	(1.295)	(1.138)
Dummy for arriving later than work restriction	-0.6271***				-0.6001***
	(-6.192)				(-5.751)
Dummy for departing earlier than home restriction	-0.3155***				-0.1402
	(-4.425)				(-1.497)
Dummy for departing later than home restriction	-0.1782*				-0.2735**
	(-1.164)				(-2.309)
Dummy for departing 30-59 min earlier than PDT	0.2963***	0.2923**	0.3017***	0.2930***	0.2934***
	(8.968)	(8.868)	(9.132)	(8.902)	(8.837)
Dummy for departing 60-89 min earlier than PDT	-0.1356***	-0.1379***	-0.1374***	-0.1376***	-0.1331***
	(-3.637)	(-3.701)	(-3.680)	(-3.692)	(-3.565)
Dummy for departing more than 90 min PDT	-0.5274***	-0.5192***	-0.5347***	-0.5198***	-0.5209***
	(-6.012)	(-5.930)	(-6.085)	(-5.936)	(-5.918)
Log-likelihood	-15387.63	-15414.93	-15399.98	-15413.79	-15367.50
Pseudo-R-sqrd	0.09474	0.09313	0.0940	0.09320	0.09575

Notes: t-statistics are shown in parentheses. Significance is indicated by ***, **, and *, referring to significance at the 99%, 95%, and 90% level, respectively.

Appendix 4E: Estimation results of trip length, income, gender, education, and cost compensation effects

Explanatory variables	Model 11		Model 12		Model 13		Model 14		Model 15	
	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.
ASC of public transport alternative	-0.8396***	(-15.381)	-0.7713***	(-14.834)	-0.7674***	(-14.731)	-0.7690***	(-14.782)	-0.8164***	(-15.565)
Travel cost C	-0.1058***	(-16.535)	-0.0971***	(-16.987)	-0.0919***	(-16.120)	-0.0941***	(-16.543)	-0.1098***	(-18.632)
E[T]	-0.0117***	(-5.997)	-0.0086***	(-4.712)	-0.0130***	(-8.698)	-0.0140***	(-9.078)	0.0037	(1.583)
E[T]*Trip length L2 (30-60 km)	0.0002	(0.078)								
E[T]*Trip length L3 (>60 km)	-0.0078***	(-3.788)								
E[T]*Income2 (household inc. €28,500-45,000)			-0.0019	(-1.075)						
E[T]*Income3 (household inc. €45,000-68,000)			-0.0120***	(-6.573)						
E[T]*Income4 (household inc. >€68,000)			-0.0084***	(-4.947)						
E[T]*Female					-0.0034**	(-2.036)				
E[T]*Higher education (HBO and above)							-0.0001	(-0.050)		
E[T]*Full compensation of travel cost									-0.0178***	(-7.616)
E[T]*Partial compensation of travel cost									-0.0229***	(-9.932)
E[SDE]	-0.0262***	(-9.286)	-0.0231***	(-7.877)	-0.0159***	(-6.902)	-0.0176***	(-7.172)	-0.0201***	(-5.169)
E[SDE]*Trip length L2 (30-60 km)	0.0094***	(3.215)								
E[SDE]*Trip length L3 (>60 km)	0.0066*	(1.833)								
E[SDE]*Income2 (household inc. €28,500-45,000)			0.0049	(1.492)						
E[SDE]*Income3 (household inc. €45,000-68,000)			0.0074**	(2.177)						
E[SDE]*Income4 (household inc. >€68,000)			0.0031	(1.239)						
E[SDE]*Female					-0.0107***	(-3.508)				
E[SDE]*Higher education (HBO and above)							-0.0032	(-1.251)		
E[SDE]*Full compensation of travel cost									0.0045	(1.139)
E[SDE]*Partial compensation of travel cost									-0.0037	(-0.921)
E[SDL]	-0.0345***	(-9.490)	0.0299***	(-8.022)	-0.0230***	(-8.981)	-0.0266***	(-9.295)	-0.0288***	(-5.803)
E[SDL]*Trip length L2 (30-60 km)	0.0153***	(3.837)								
E[SDL]*Trip length L3 (>60 km)	0.0185***	(4.359)								
E[SDL]*Income2 (household inc. €28,500-45,000)			0.0032	(0.733)						
E[SDL]*Income3 (household inc. €45,000-68,000)			0.0133***	(2.948)						
E[SDL]*Income4 (household inc. >€68,000)			0.0109***	(2.718)						
E[SDL]*Female					0.0010	(0.262)				
E[SDL]*Higher education (HBO and above)							0.0077**	(2.440)		
E[SDL]*Full compensation of travel cost									0.0089*	(1.703)
E[SDL]*Partial compensation of travel cost									0.0049	(0.952)
Probability of late arrival (later than PAT)	-0.0850*	(-1.734)	-0.0932*	(-1.915)	-0.0893*	(-1.836)	-0.0906*	(-1.864)	-0.099**	(-2.023)
Uncertainty	0.0010	(0.707)	0.0014	(1.472)	0.0018	(1.284)	0.0020	(1.405)	0.0015	(1.042)
Dummy for departing 30-59 min earlier than PDT	0.2302***	(6.475)	0.2902***	(8.799)	0.3112***	(9.384)	0.2913**	(8.841)	0.2644***	(7.937)
Dummy for departing 60-89 min earlier than PDT	-0.1359***	(-3.625)	-0.1405***	(-3.763)	-0.1418***	(-3.808)	-0.1366***	(-3.660)	-0.1508***	(-4.032)
Dummy for departing more than 90 min PDT	-0.3919***	(-4.358)	-0.5085***	(-5.797)	-0.5489***	(-6.249)	-0.5187	(-5.919)	-0.4631***	(-5.266)
Log likelihood	-15390.69		-15384.71		-15405.94		-15413.49		-15355.98	
Pseudo-R-sqrd	0.09450		0.09479		0.09367		0.09322		0.09655	

Notes: t-statistics are shown in parentheses. Significance is indicated by ***, **, and *, referring to significance at the 99%, 95%, and 90% level, respectively.

Appendix 4F: Explanation and population share of explanatory (dummy) variables of data set (N=1115)

Categories	Definitions and population share
Gender	Male = 1 if male (76.23%) Female = 1 if female (23.77%)
Education	Lower education = 1 if senior general secondary (HAVO/VWO) or lower (55.25%) Higher education = 1 if Bachelor (HBO/WO) or higher (44.75)
Income	Income 1 = 1 if household gross yearly income is less than 28,500 euros (20.72%) Income 2 = 1 if household gross yearly income is 28,500 – 45,000 euros (26.73%) Income 3 = 1 if household gross yearly income is 45,000 – 68,000 euros (26.10%) Income 4 = 1 if household gross yearly income is more than 68,000 euros (26.46%)
Trip length	Trip length L1 = 1 if the usual commuting distance is less than 30 km (35.16%) Trip length L2 = 1 if the usual commuting distance is 30 - 60 km (36.95%) Trip length L3 = 1 if the usual commuting distance is more than 60 km (27.89%)
Late arrival time restriction	Late arrival time restriction =1 if commuters cannot arrive at work later than certain time (54.71%)
Early departure time restriction	Early departure time restriction =1 if commuters cannot depart from home earlier than certain time (15.07%)
Late departure time restriction	Late departure time restriction = 1 if commuters cannot depart from home later than certain time (14.44%)

CHAPTER 5

5 VALUE OF TIME BY TIME OF DAY: A STATED PREFERENCE STUDY

5.1 INTRODUCTION

The value of travel time savings (VTTS), often abbreviated as the ‘value of time’ (VOT), is of central interest in transport research. The VTTS is often one of the bigger benefit components in the assessment of transport investments. It is also an important parameter in the analysis of travel behavior and in traffic assignment models. Becker (1965) was probably the first to introduce the allocation of time over various activities in the analysis of consumer behavior, thus offering the microeconomic framework needed to establish the shadow price of time savings. Further contributions from Johnson (1966) who introduced work time in the utility function, Oort (1969) who did the same for travel time, and DeSerpa (1971) who added technical constraints, showed that the VOT for various activities does not need to be equal to the wage rate – justifying further research into the empirical estimation of the VTTS. In such research, the VTTS is usually derived as the marginal rate of substitution between travel time and cost coefficients, typically as found in discrete choice models of stated preference data, revealed preference data, or a combination of these (e.g. Small et al. 2005). This ratio is exactly the VTTS in DeSerpa’s (1971) framework (Bates, 1987).

An important addition to this framework was made by Small (1982), who explicitly included the scheduling of activities – the morning commute in particular – in the analysis. Inspired by the work of Vickrey (1969) on the dynamic equilibrium and optimum for queuing behind a bottleneck, Small allowed for disutility from early or late arrivals at work. With a simple linear utility specification, this leads to three relevant time-related shadow prices: a value of travel delay which is usually denoted α in the relevant literature (e.g. Arnott et al., 1993), a shadow price of arriving early (β), and a shadow price of late arrivals (γ). Small’s (1982) model has become the workhorse model to incorporate within-the-day dynamics in the valuation of travel time components.

A particular aspect of Small's basic linear model is that the inconvenience of an early trip is attributed to an early arrival at work. This implies, in terms of the notation just introduced, that a constraint $\beta<\alpha$ should be imposed – unless one is willing to accept that an individual may prefer to stay in the parked car or to keep on driving around the block when arriving early, over getting out of the car and into the office (or factory). Moreover, $\beta>\alpha$ also implies that, for a deterministic dynamic equilibrium with homogeneous travelers, a person who arrives after another person should have departed before that other person for the total trip cost to be constant by arrival time (which is the natural dynamic equilibrium condition for homogeneous travelers). But consistent overtaking as a structural equilibrium phenomenon is not plausible.

Nevertheless, the inequality $\beta<\alpha$ need not always be satisfied in applied work, especially not for linear specifications. Indeed, the focus on deviations from desired arrival times, in combination with the constancy of shadow prices α, β and γ for a linear utility function, makes it impossible to capture the intuitive notion that rescheduling to earlier time slots becomes increasingly inconvenient as increasingly more valuable moments spent at home (possibly sleeping) are to be given up. However, allowing for this by introducing a nonlinear effect of schedule delay early upon utility is somewhat ad hoc: it can partly account for the effect mentioned, but leads to the risk of $\beta>\alpha$ for early arrivals and may assign different penalties for early departures from home at the same time when trip durations vary.

Our purpose is to propose a variant of Small's model that explicitly allows for nonlinear effects. We present a model that estimates two time-dependent utility components: namely, the per-unit-of-time willingness-to-pay (WTP) for being at home over being in the car $H(t)$ (a negative value would imply a preference for the car), and the per-unit-of-time WTP for being at work over being in the car $W(t)$. We show how the implied two functions in fact define three time-dependent functions $\alpha(t)$, $\beta(t)$ and $\gamma(t)$, which are the time-varying equivalents of the conventional shadow prices α, β and γ. A recent study of Liu et al. (2007) provides empirical evidence that the values of time (VOT) and unreliability (VOR) vary over the peak. Our approach differs from theirs in at least two respects: first, we explicitly include scheduling considerations in the model; and secondly, our approach allows an individual's VOT to vary over time, whereas the empirical time-dependency of the VOT in the analysis of Liu et al. (2007) may equally well result from different departure time choices from different travelers with different VOTs and VORs.

Section 5.2 presents our framework and discusses how it has the conventional linearized version of Small's (1982) model as a special, restricted case. Section 5.3 introduces the empirical application of our model, and Section 5.4 presents the estimation results. Section 5.5 concludes.

5.2 THEORETICAL FRAMEWORK

Consider an individual who has to decide whether or not to travel to work in the morning, and, if so, at what time. We hypothesize that this individual's utility over the full morning period (that is, a period between two instants t_b ("begin") and t_e ("end"), chosen such that it is long enough to encompass all possibly relevant departure times from home and arrival times at work) can be found by integrating, over the relevant periods, particular functions that represent the per-unit-of-time utility of being at a certain place. To simplify matters, let us assume that the individual considers three possible places where he can be: at home, in the vehicle, or at work (all other possible places are considered inferior to even the least preferable of these three at any relevant instant t). Let every unit of time spent at home produce a time-varying utility $h(t)$, and define analogously the utility in the vehicle $v(t)$, and at work $w(t)$. Assume that each of these functions is continuous and smooth.

We ignore complexities that may stem from the dependence of the marginal utility of income upon the scheduling decisions in the morning period, and assume it is constant. We can then choose units of utility such that differences between any pair from the triplet $\{h(t), v(t), w(t)\}$ denote the individual's WTP to spend, at time t, one unit of time at the more-preferred location rather than at the less-preferred one.

Indeed, because we only consider three possible locations where the individual can be, behaviour will be determined entirely by differences in utility levels. We can therefore simplify the notation by equating one of the three utility levels to 0 throughout the period $[t_b, t_e]$, which only means that calculated utility levels are reduced by a constant equal to the integral of this reference utility function between t_b and t_e. We choose $v(t)$ as the reference, so that we are left with two functions that we define as follows:

$$\begin{aligned} H(t) &= h(t) - v(t); \\ W(t) &= w(t) - v(t). \end{aligned} \tag{5.1}$$

H(t) therefore gives the WTP to spend a unit of time at home rather than in the vehicle at time *t*, and *W(t)* does the same for work versus the vehicle. We will refer to the functions *H(t)* and *W(t)* as excess-willingness-to-pay functions (hereafter EWTP-functions). Note that there is no a priori reason for restricting either EWTP to be positive; a negative value would be consistent with findings reported by, among others, Redmond and Mokhtarian (2001), on how travelers may sometimes attach a positive utility to extra time spent in the car.

Figure 5.1 provides an example, where we made the plausible assumptions that the individual finds it, especially at early hours, attractive to spend time at home rather than in the vehicle, while the attractiveness of being at work increases rapidly within a relatively short time-span around the official work start time, and remains rather flat both before and after that moment.

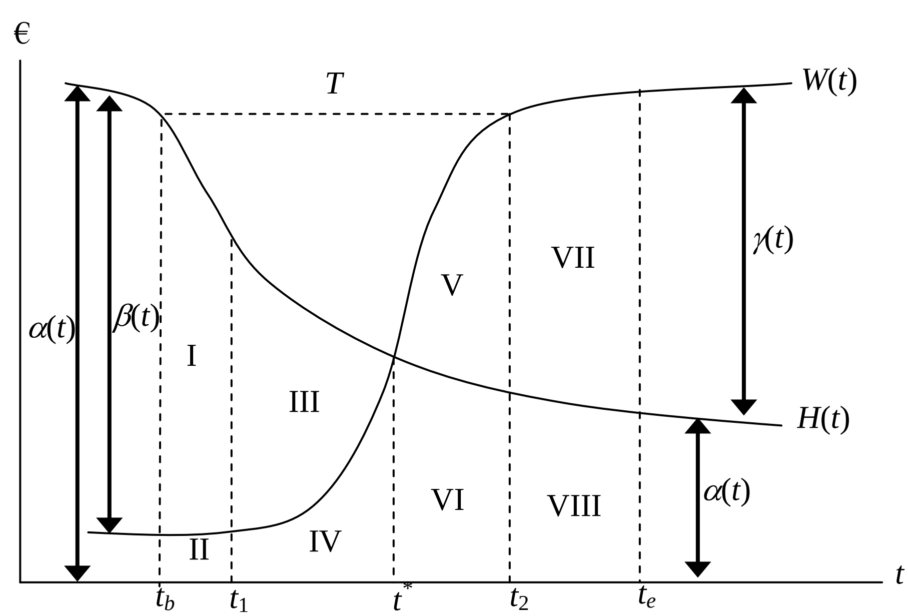


Figure 5.1 The EWTP-functions for being at home (H(t)) and at work (W(t))

Let us now define t_D and t_A as the moments of departure and arrival. If the trip from home to work were to require no travel time, so that $t_A = t_D$, the individual would travel at the instant that $H(t)$ and $W(t)$ intersect, at t^*. It is tempting to call t^* the "desired arrival time", but it actually is most desirable only if travel takes no time. With a *given* trip duration T, the individual would otherwise prefer to depart and arrive at moments t_D and t_A, such that $H(t_D)$ and $W(t_A)$ are equalized while $t_A - t_D = T$ is satisfied. The trip between t_b and t_2 in Figure 5.1 is an example of such an optimally-timed trip, for which t_2 is the desired arrival time[19]. Therefore, because t^* is the desired arrival time only when travel time is zero, we will refer to t^* as the "ideal arrival time".

It seems natural to define the individual's travel cost $c(t_D,t_A)$ such that it is zero for an optimally timed zero-duration trip: $c(t^*,t^*) = 0$. The WTP for being able to make this trip, over the worst possible situation of being in the vehicle between instants t_b and t_e, is given by the sum of all areas I – VIII in Figure 5.1. Travel cost $c(\cdot)$ for other trips can be identified graphically in Figure 5.1 as the areas below the maximum of $H(t)$ and $W(t)$ when driving, and between $H(t)$ and $W(t)$ between moments t_A and t^* when arriving before t^*, and between t^* and t_D when departing after t^*. These areas together thus give the WTP for making the ideal optimally-timed zero-duration trip, over the trip under consideration.

Another way to identify the same areas for a given trip is to take the area below $H(t)$ when driving (*i.e.*, between t_D and t_A), and add to it the area between $H(t)$ and $W(t)$ between moments t_A and t^* when arriving before t^*, or between t^* and t_A when departing after t^*:[20]

---

[19] With time-varying travel times, equalization of $H(t_D)$ and $W(t_A)$ is, of course, no longer the appropriate necessary equilibrium condition. The optimality condition for such cases is straightforward to express after specifying a travel time function $T(t_D)$ and replacing t_A by $t_D+T(t_D)$. It next involves minimization of equation (2) below with respect to t_D. This leads to the equilibrium condition: $H(t_D) = [1+\mathrm{d}T(t_D)/\mathrm{d}\ t_D]\cdot W(t_A)$. Note how it has the constant travel time case discussed in the main text as a special case, where the second term in the square bracket is zero. Otherwise, it corrects for changes in travel time where the departure time is marginally adjusted. With time-varying travel times, an arrival before t^* or a departure after t^* may occur in equilibrium.

[20] Equivalently, we could specify the model such that $W(t)$ becomes the time-varying value of travel time. This involves schedule delay terms that are the integral of $[H(t)–W(t)]$ between t_D and t^* for departures before t^*, and the integral of $[W(t)–H(t)]$ between $t^*$and t_D for departures after t^*. The fact that we can do this is consistent with our model being symmetric between $H(t)$ and $W(t)$. The specification in the main text has the obvious advantage of being directly comparable to the conventional model. The alternative described in this footnote, which relates schedule delay costs to deviations from the ideal *departure time*, might be more intuitive to describe the afternoon peak.

$$c(t_D, t_A) = \int_{t_D}^{t_A} H(t)\mathrm{d}t + \begin{cases} \int_{t_A}^{t^*} \left(H(t) - W(t)\right)\mathrm{d}t & \text{if} \quad t_A < t^* \\ \int_{t^*}^{t_A} \left(W(t) - H(t)\right)\mathrm{d}t & \text{if} \quad t_A \geq t^* \end{cases} . \tag{5.2}$$

This can be verified by checking the following travel costs for three different types of trips in Figure 5.1: namely, one with both the departure and arrival before t^*, one with a departure before t^* and an arrival after t^*, and one with both departure and arrival after t^*:

$$\begin{aligned} c(t_b, t_1) &= \mathrm{I} + \mathrm{II} + \mathrm{III} \\ c(t_1, t_2) &= \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI} \\ c(t_2, t_e) &= \mathrm{V} + \mathrm{VII} + \mathrm{VIII} \end{aligned} \tag{5.3}$$

It is now convenient to define the following functions:

$$\begin{aligned} \alpha(t) &= H(t) \\ \beta(t) &= H(t) - W(t) \\ \gamma(t) &= W(t) - H(t) \end{aligned} . \tag{5.4}$$

We can then rewrite the cost function in (5.2) as:

$$c(t_D, t_A) = \int_{t_D}^{t_A} \alpha(t)\mathrm{d}t + \begin{cases} \int_{t_A}^{t^*} \beta(t)\mathrm{d}t & \text{if} \quad t_A < t^* \\ \int_{t^*}^{t_A} \gamma(t)\mathrm{d}t & \text{if} \quad t_A \geq t^* \end{cases} . \tag{5.5}$$

The arrows in Figure 5.1 represent these functions $\alpha(t)$, $\beta(t)$ and $\gamma(t)$. These functions are helpful in understanding how the conventional linearized version of Small's (1982) model is a special case of our model. The cost function in the conventional linear model can be written as follows:

$$c(t_D, t_A) = \alpha \cdot \left(t_A - t_D\right) + \begin{cases} \beta \cdot \left(t^* - t_A\right) & \text{if} \quad t_A < t^* \\ \gamma \cdot \left(t_A - t^*\right) & \text{if} \quad t_A \geq t^* \end{cases} . \tag{5.6}$$

This is, of course, a special case of the cost function in (5.5), with the functions $\alpha(t)$, $\beta(t)$ and $\gamma(t)$ all constant over time and equal to α, β and γ, respectively. Figure 5.2 depicts the associated variant of Figure 5.1.

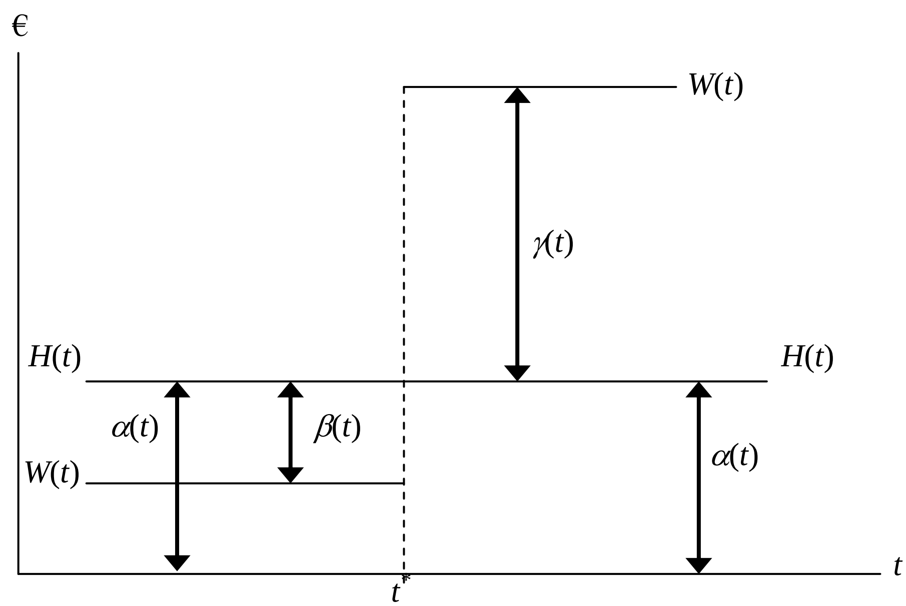


Figure 5.2 The EWTP-functions in the conventional linearized model

The most striking difference between Figures 5.1 and 5.2 is that $H(t)$ is constrained to be constant in the latter, because $\alpha(t)$ is. The inconvenience of an early schedule insofar as affecting the time otherwise spent home, perhaps sleeping, can therefore not be properly reflected. A minute spent at home, instead of in the vehicle, is constrained to be just as valuable at 5 a.m. as at 8 a.m. Introspection suggests that this is not plausible, and empirical evidence in the next section will confirm this. The other difference is that $W(t)$ is piecewise constant in Figure 5.2. Although this is restrictive too, it seems less so than the first difference – of imposing constancy of $H(t)$ over time.

Finally, it can be noted that the proposed framework is related to recent developments in activity-based modelling (e.g. Ettema and Timmermans, 2003; Ashiru et al., 2004), which also involve the modelling of dynamic scheduling decisions. Our approach is different in that we characterize time-dependent utility functions entirely in terms of time-varying shadow prices, the

EWTP-functions, as a dynamic generalization of the conventional constant shadow prices α, β and γ.

5.3 EMPIRICAL APPLICATION

5.3.1 Empirical specifications

The conventional linearized model of equation (5.6) is easily operationalized for application in a random-utility discrete-choice model through the following specification of a linear systematic utility function:

$$V = \beta_\alpha \cdot T + \beta_\tau \cdot \tau + \begin{cases} \beta_\beta \cdot (t^* - t_A) & \text{if} \quad t_A < t^* \\ \beta_\gamma \cdot \left(t_A - t^*\right) & \text{if} \quad t_A \geq t^* \end{cases}, \tag{5.7}$$

with τ denoting a monetary attribute such as a toll. With the coefficients β estimated, one can immediately determine $\alpha = \beta_\alpha/\beta_\tau$; $\beta = \beta_\beta/\beta_\tau$; and $\gamma = \beta_\gamma/\beta_\tau$. With random utility for alternative j for individual n defined as:

$$U_{jn} = V(\cdot) + \varepsilon_{jn}, \tag{5.8}$$

conventional multinomial logit or probit discrete-choice models arise under particular assumptions on the distribution of the random terms ε_{jn}.

The operationalization of the more flexible model of (5.2) and (5.4) is less straightforward. One option would be to impose functional forms for $H(t)$ and $W(t)$ as functions of time, and to estimate the relevant parameters. Another possibility, pursued below, could be characterized as 'flexible' and divides the morning peak period up into a number of smaller intervals, with $H(t)$ and $W(t)$ constant within an interval but free to vary between them. This has the advantage of not imposing any a priori assumption on the possible time patterns of these functions, but the obvious disadvantage of restricting $H(t)$ and $W(t)$ to be constant within an interval. Because we are interested primarily in detecting the pattern of time variation, if any, in $\alpha(t)$ over the peak, we judged the advantage to outweigh the disadvantage.

The systematic utility function can be written after defining T_i as the amount of time spent driving during time interval *i*, T_i^E as the amount of time spent at work before the ideal arrival time t^* during time interval *i*, and T_i^L as the amount of time not (yet) spent at work after t^* during time interval *i*. This leads to the following discrete-time version of equation (5.5):

$$c(t_D, t_A) = \sum_i \alpha_i \cdot T_i(t_D, t_A) + \sum_i \beta_i \cdot T_i^E(t_A) + \sum_i \gamma_i \cdot T_i^L(t_A) \quad . \tag{5.9}$$

The accompanying systematic utility function for estimation purposes becomes:

$$V = \beta_\tau \cdot \tau + \sum_i \beta_{\alpha,i} \cdot T_i + \sum_i \beta_{\beta,i} \cdot T_i^E + \sum_i \beta_{\gamma,i} \cdot T_i^L \quad . \tag{5.10}$$

After estimating the coefficients β, one can immediately determine $\alpha_i = \beta_{\alpha,i}/\beta_\tau$; $\beta_i = \beta_{\beta,i}/\beta_\tau$; and $\gamma_i = \beta_{\gamma,i}/\beta_\tau$. As a final step, one could next calculate $H_i = \alpha_i$, and $W_i = \alpha_i + \beta_i$ for early arrivals and $W_i = \alpha_i + \gamma_i$ for late arrivals. Note that we assume intervals to be defined such that t^* defines the boundary between two intervals.

5.3.2 Data

The data used in this study were obtained from the second part of choice experiment of MD-PIT project. The details of the survey and data collection of MD-PIT project were described in Section 4.1. Table 4.2 illustrated the design of the stated choice experiment, and Appendix 4B showed an example of the choice screen presented to the respondents.

The dependent variable used in the present study is the choice proportion allocated to each of the four alternatives by individuals. Uncertainty in travel time is a separate attribute in this stated choice experiment, and we follow the 'expected utility maximization' approach of Noland and Small (1995) to incorporate the uncertainty variable. This means that we compute the 'expected' travel time and early/late arrival in different time intervals. In what follows, T_i, T_i^E and T_i^L thus refer to expected numbers of minutes in time interval *i*.

The choice experiment was not constructed specifically for estimating the above-specified time-dependent model, which caused restrictions in the estimation of equation (5.10). A first restriction is that an individual's preferred arrival time is available only for a trip with a duration implied by the individual's free-flow (uncongested) travel time. This preferred arrival time *PAT*, and the accompanying preferred departure time *PDT*, define an interval in which the ideal arrival time t^* should be located (see Figure 5.1). We will not attempt to actually estimate t^*, but instead test a few specifications in which t^* is defined as a weighted average, with weights equal across individuals, of *PDT* and *PAT*. We tested various weights between 0 and 1, and the results appeared to be robust for various weights.

A second restriction is that the design of the experiment does not provide sufficient variation in arrival times at work to allow for unrestricted estimation of all coefficients β for all time periods of interest. Since there are hardly any observations of a departure time after t^*, there is a nearly perfect correlation between T_i^L and T_i in intervals after t^*. We can therefore only estimate the sum of α_i and γ_i for these late intervals. To ensure that there are sufficient observations in each time interval, we chose 30 minutes as the typical interval size. Still, since most arrivals are less than 30 minutes from t*, we were able to use two smaller intervals of 15 minutes just before and after t*. The notation, description and descriptive statistics of the variables are shown in Table 5.1. The utility function we actually estimate is then given by:

$$V = \beta_\tau \cdot \tau + \sum_{i=-7}^{-1} \beta_{\alpha,i} \cdot T_i + \sum_{i=-2}^{-1} \beta_{\beta,i} \cdot T_i^E + \sum_{i=1}^{3} \left(\beta_{\alpha,i} + \beta_{\gamma,i}\right) \cdot T_i^L \ . \tag{5.11}$$

Table 5.1 The definition and descriptive statistics of the explanatory variables

Variable notation	Definition	Mean	Std.	Min.	Max.
T_{-1}	The expected *travel time* spent during the interval between **0-15** minutes before t*	10.19	6.28	0	15
T_{-2}	The expected *travel time* spent during the interval between **15-30** minutes before t*	12.20	4.57	0	15
T_{-3}	The expected *travel time* spent during the interval between **30-60** minutes before t*	24.12	10.34	0	30
T_{-4}	The expected *travel time* spent during the interval between **60-90** minutes before t*	16.03	13.41	0	30
T_{-5}	The expected *travel time* spent during the interval between **90-120** minutes before t*	8.18	12.09	0	30
T_{-6}	The expected *travel time* spent during the interval between **120-150** minutes before t*	3.62	8.96	0	30
T_{-7}	The expected *travel time* during the interval between **150-180** minutes before t*	1.52	6.07	0	30
T^E_{-1}	The expected *time spent at the destination* during the interval between **0-15** minutes before t*	4.57	6.24	0	15
T^E_{-2}	The expected *time spent at the destination* during the interval between **15-30** minutes before t*	1.95	3.81	0	14
T^L_1	The expected *time not spent at the destination* during the interval between **0-15** minutes after t*	7.56	7.30	0	15
T^L_2	The expected *time not spent at the destination* during the interval between **15-30** minutes after t*	4.45	6.46	0	15
T^L_3	The expected *time not spent at the destination* during the interval between **30-60** minutes after t*	1.85	3.70	0	27
Cost	Sum of fuel cost and toll (in euros)	3.54	3.67	0	33.7
Uncertainty	The difference between maximum and minimum possible travel times (in minutes)	15.93	16.43	0	168

5.4 ESTIMATION RESULTS

The choice models we estimate in this section are based on the specification of equation (5.11). The objective is to see if the estimated coefficients vary over the different time intervals, and, more importantly, whether the variation of the estimates follows the pattern illustrated in Figure 5.1. Since the experiment involves mode choice, we specify an alternative-specific constant (ASC_{PT}) for public transport, in order to capture the effect of respondents' preference associated with that particular mode. Furthermore, we add the variable 'uncertainty' for car alternatives, defined as the width of the possible arrival time interval as shown to the respondents, in order to account for the additional disutility associated with travel time uncertainty (apart from the scheduling costs). The public transport alternative was presented as completely reliable in the experiment.

As discussed in the previous section, we should expect an individual's t^* to be somewhere between the individual's preferred departure time (*PDT*) and preferred arrival time (*PAT*) for a free-flow travel time trip. Because *PDT* and *PAT* were asked in the questionnaire and t^* was not, the latter should somehow be derived from the former two. We have estimated a series of models by varying the individuals' locations of t^*, relative to their *PDT* and *PAT*. The best model, pragmatically defined as the one yielding the highest log-likelihood value, is the one with $t^*=0.2*PDT + 0.8*PAT$. The choice models that we summarize in Table 5.2 are based on that particular specification.

The estimation results are summarized in Table 5.2. Two models are presented: (1) a conventional multinomial logit (MNL) model, and (2) a mixed logit (ML) model that accommodates the correlation between choice sets drawn from the same individual. For model stability, cost parameters were treated as non-random in our ML models (see also Revelt and Train, 1998; Bhat and Sardesai, 2006). The random parameters in the mixed logit model are assumed to follow a normal distribution[21].

The main result is that the values of travel time savings α_i are indeed not constant over time, and the values increase as the time intervals move further away from t^*. These results imply that individuals do have shadow prices that vary over time. The implied marginal rates of substitution between time attributes and money, i.e. the various shadow prices of interest, are shown in Table 5.3, and the patterns generally follow those hypothesized in Figure 5.1.

The design of the underlying questionnaire in the first place allows us to provide estimates of α_i, and hence the function H, for all intervals before t^*. For example the MNL estimates of Model 1 depict how this value dramatically falls over time during the period prior to t^*, with the time coefficient becoming insignificantly different from zero for intervals –2 and –1, reflecting that the individuals are indifferent between spending those minutes at home or in the car. This pattern thus replicates the hypothesized pattern in Figure 5.1, and suggests that the usual assumption that α is constant over time may be rather restrictive. We can estimate β_i only for intervals –1 and –2,

[21] A number of ML models with uniform and triangular distributions were also estimated, and the results were similar to the models with a normal distribution. The models with normal distribution yield higher log-likelihood values. We also estimated the models with log-normal distribution, but most of these hardly converged.

and therefore the same holds for the determination of the function W before t^*. For both models, the associated time coefficient is significantly different from zero only in interval –2. Because β_{-2} exceeds α_{-2}, it appears that W_{-2} is negative: during that time interval, people prefer to be in the car over already being at the destination. For intervals after t^*, we can only estimate the sum $\alpha_i+\gamma_i$, which is equal to W_i, and not the two components separately. (Because no separate estimate of α_i can be made, also H_i can not be determined for intervals after t^*.) The pattern clearly shows a sharp increase from interval 1 to interval 3, which again confirms our hypotheses illustrated in Figure 5.1.

Table 5.2 Estimation results

Explanatory variable	Refers to valuation	Model 1 MNL Coeff.	Model 1 MNL t-stat.	Model 2 ML Coeff.	Model 2 ML t-stat.
Random parameter mean effects					
T_{-1}	α_{-1}	.0040	0.36	.0082	0.55
T_{-2}	α_{-2}	.0002	0.05	.0145	2.01
T_{-3}	α_{-3}	-.0107	-6.19	-.0172	-4.26
T_{-4}	α_{-4}	-.0215	-11.63	-.0416	-10.89
T_{-5}	α_{-5}	-.0269	-10.24	-.0471	-9.78
T_{-6}	α_{-6}	-.0311	-7.03	-.0544	-6.28
T_{-7}	α_{-7}	-.0414	-6.22	-.0592	-3.69
T^E_{-1}	β_{-1}	.0063	0.53	.0137	0.81
T^E_{-2}	β_{-2}	-.0370	-5.14	-.0244	-1.92
T^L_1	$\alpha_1+\gamma_1$	-.0173	-4.42	-.0220	-3.51
T^L_2	$\alpha_2+\gamma_2$	-.0370	-6.76	-.0456	-4.42
T^L_3	$\alpha_3+\gamma_3$	-.0299	-3.88	-.0576	-3.91
Uncertainty		-.0027	-2.00	-.0079	-2.32
Random parameter standard deviation					
T_{-1}				.0075	0.64
T_{-2}				.0086	13.91
T_{-3}				.0087	23.79
T_{-4}				.0064	19.45
T_{-5}				.0046	10.17
T_{-6}				.0048	7.35
T_{-7}				.0095	6.40
T^E_{-1}				.0032	0.35
T^E_{-2}				.0043	0.21
T^L_1				.0022	2.92
T^L_2				.0088	12.35
T^L_3				.0058	3.84
Uncertainty				.0060	28.29
Non-random parameters					
Cost		-.0989	-17.18	-.1957	-34.77
ASC, PT		-.9496	-18.34	-1.487	-24.86
Observations (N)		12265		12265	
Log likelihood		-15401.38		-13141.79	
Pseudo-R-sqrd		0.0921		0.2265	

Table 5.3 Mean monetary values (€/hour) (standard deviations of distributions of random parameters)

Monetary values	Model 1 MNL	Model 2 ML
α_{-1}	*-2.43*	*-2.50* (*2.31*)
α_{-2}	*-0.13*	-4.46 (26.22)
α_{-3}	6.48	5.28 (26.54)
α_{-4}	13.03	12.75 (19.76)
α_{-5}	16.35	14.44 (14.17)
α_{-6}	18.87	16.69 (14.71)
α_{-7}	25.15	18.16 (29.05)
β_{-1}	*-3.84*	*-4.21* (*0.97*)
β_{-2}	22.46	7.49 (*1.31*)
$\alpha_1+\gamma_1$	10.47	6.75 (6.72)
$\alpha_2+\gamma_2$	22.47	13.97 (26.97)
$\alpha_3+\gamma_3$	18.12	17.67 (17.93)
VUNC (uncertainty)	1.65	2.43 (18.54)

Notes: 1. The derived monetary values from the standard deviation of random parameters are shown in parentheses.
2. Values or standard deviations arising from insignificant (at the 90% significance level) coefficients are in italics.

Model 2 is preferred to Model 1 since the former allows for taste heterogeneity across individuals, as well as accommodating the correlations across choice sets that are drawn from the same individual (ML with panel structure). The LR test also suggests that the model fit for Model 2 is better than it is for Model 1 at any level of significance. Although the mean monetary values differ somewhat between Models 1 and 2 (see Table 5.3), the two models seem to produce reasonably consistent results in terms of the qualitative patterns. It is interesting to note that the standard deviations of random parameters in Model 2 are mostly significant and large. It makes intuitive good sense that there is variation in parameters across individuals. Note that the negative ASC_{PT} in both models suggests that our respondents have an inherent preference for car over public transport. This is consistent with the fact that our sampled respondents are all frequent car users.

Figure 5.3 summarizes the above findings graphically, by comparing the estimated time-dependent patterns of H and W from the two time-dependent models (1 and 2) with the patterns that result from the conventional linear model estimated for the same data, for which we used estimates of α, β and γ as reported in Model 2 of Chapter 4 (see Table 4.9). The horizontal axis depicts time, where the ideal arrival time t^* is set at t=0 for the time-dependent model, and the preferred arrival time *PAT* for the linear model. The falling pattern of *H* for times before t^* is, by construction, not found in the linear model, but the time-independent value of *H* seems to be reasonably close to the time average for the time-dependent models. The negative values for *W* in interval –2 in the time-dependent models are consistent with a consistently negative value of *W* before *PAT* in the linear model ($\beta > \alpha$ in that model). Apparently, there is a tendency among the respondents to be keen to avoid spending time at work before the start time: time spent in the car is valued higher than time spent at work in our estimates. For times after t^*, the linear model implies that *H* maintains its pre-t^* value; there are insufficient observations to estimate the corresponding values for the time-dependent model, but the last value estimated, not significantly different from zero, suggests that lower values would have resulted, just as hypothesized in Figure 5.1. Finally, *W* rises as the time interval progresses, but stays below the value implied by the linear model. The biggest difference between the time-dependent models versus the linear model seems, therefore, to be that the latter cannot reproduce the falling pattern of *H*, corresponding to the conventional value of travel time savings α, over the period prior to the ideal arrival time t^*. Comparing the two time-dependent models, the various estimates seem reasonably close, with the ML model producing less extreme values and smoother patterns than its MNL counterpart.

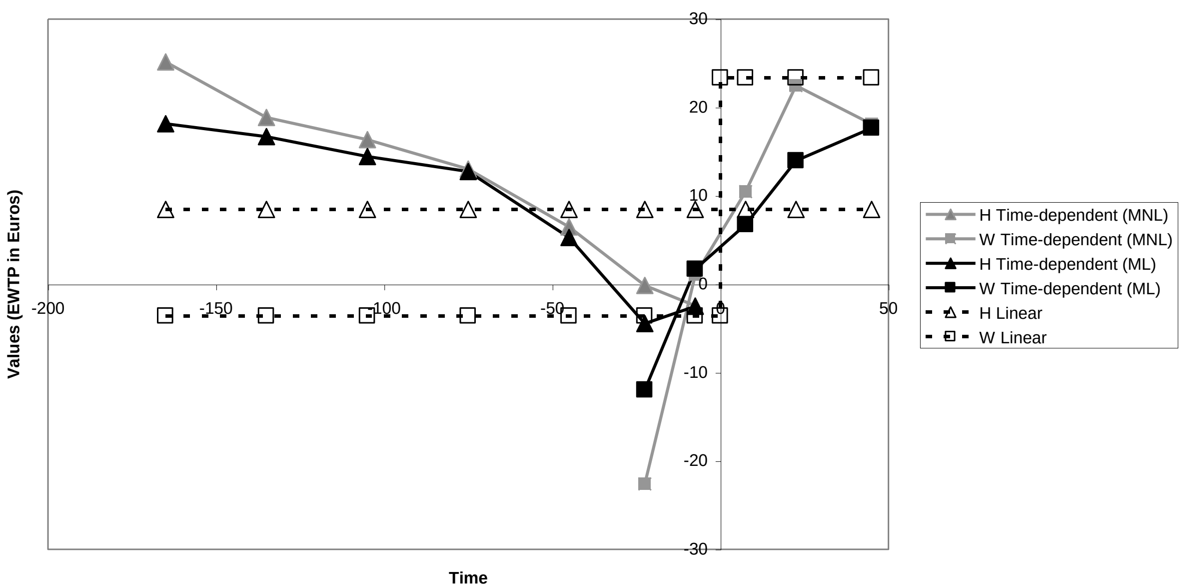


Figure 5.3 The empirical EWTP-functions for the time-dependent and the linear models

There is one final potential caveat to be addressed. Because the design of the SCE was such that arrival times were varied around the most desired arrival time, we have relatively many observations for respondents making longer trips in the earlier intervals. Now, if for some reason these individuals have a higher value of travel time savings, the patterns depicted for *H* in Figure 5.3 might also be caused by a different sampling of respondents along the time axis. We have therefore also estimated the same time-dependent models for the 48 percent of respondents who make the longest trips (>40 km). The results conveyed the same qualitative patterns as those shown in Figure 5.3[22].

We also verified whether the results vary by socioeconomic characteristics. Estimation results, reworked into monetary values as in Table 5.3, are provided in the Appendix 5A. A separation by income into two groups left the qualitative patterns, as shown in Figure 5.3, intact. It further revealed that the higher-income group has higher α's before t^*, whereas the β's for arrivals before t^*, and $\alpha+\gamma$ for arrivals after t^* are higher for lower-income groups. This confirms the expectation that higher income groups are willing to pay more to avoid travel time, whereas lower income groups have tighter scheduling constraints. A separation by gender produces somewhat less clear-cut results, partly because there are too few observations for female drivers in some time periods. But when they travel, women usual appear to have higher values than male drivers, for the shadow prices listed in Table 5.3. And finally, we tested whether the results are affected by the inclusion of a lateness penalty (a probability in our model, because we have an uncertain travel time), also introduced by Small (1982). The coefficient for this variable turned out to be statistically insignificant, both in MNL and in ML estimations, and the other coefficients are (not surprisingly) hardly affected.

5.5 CONCLUSION

We proposed an alternative, dynamic framework for estimating time-varying values of travel time savings and values of schedule delay. Our formulation represents time-preferences as the

[22] For this estimate, we had to merge the two 15 minutes intervals prior to t^* into one 30 minute interval. For the MNL model, the values of H were as follows: $H_{-1\&2} = -7.0$ (insign.); $H_{-3} = 1.3$ (insign.); $H_{-4} = 15.4$; $H_{-5} = 20.0$; $H_{-6} = 20.5$; $H_{-7} = 29.0$; and the values for W were as follows: $W_1 = 5.6$ (insign.); $W_2 = 25.3$; $W_3 = 27.1$. For the ML model, the values of H were as follows: $H_{-1\&2} = -1.8$ (insign.); $H_{-3} = -2.4$ (insign.); $H_{-4} = 12.9$; $H_{-5} = 20.0$; $H_{-6} = 19.5$; $H_{-7} = 23.3$; and the values for W were as follows: $W_1 = 7.2$; $W_2 = 11.1$; $W_3 = 20.0$.

time-varying excess willingness-to-pay (EWPT) to being in the one location over being elsewhere. We applied the framework to SP data representing the respondents' departure time choices for the morning commute. We showed how the conventional linear model is a special case of our model, and that the conventional model is implausible particularly in that it implicitly assumes that the WTP for spending a minute at home instead of being in the vehicle does not vary by time of day, not even for very early departures. It is especially in this respect that the estimates for the time-dependent model deviate substantially from those for the stationary model, estimated on the same data. The data thus support the case for our time-dependent framework rather convincingly.

Because the conventional linear model is a special, constrained case of the time-dependent model that we propose, it seems that there is little point in discussing which model is preferable from a theoretical or a behavioral viewpoint. That is, if the data allow, it seems preferable to estimate the time-dependent model and then decide whether imposing the conventional constraint of a constant value of travel time savings α seems justifiable. From a practical perspective, it is clear that there may be other considerations. If anything, a proper estimation of the time-dependent model requires a rather rich data set, with wide ranges of departure and arrival times, in particular if the EWPT for being at work is also to be estimated for intervals before the ideal arrival time t^*, and the EWPT for being at home is also to be estimated for intervals after t^*.

Our results suggest that individuals' time-related shadow prices vary strongly over the morning peak, and values of travel time savings are consequently strongly time-dependent. A failure to incorporate such considerations may produce biased estimates of values of travel time savings, and errors in the prediction of behavioral responses to policies or other measures that affect the time pattern of congestion in the morning peak. This, in turn, may of course also affect the accuracy of cost-benefit calculations for such measures. It seems difficult to predict, in general, the relative size and sign of such biases. Studying this question in the context of a dynamic equilibrium model would be an interesting topic for further study.

Appendix 5A. Monetary values in Table 5.3 disaggregated by income and gender

Monetary values by income and gender (MNL)

Monetary values	Income		Gender	
	Low[1]	High	Male	Female
	MNL	MNL	MNL	MNL
α_{-1}	*-9.29*	*4.46*	*-4.11*	*2.39*
α_{-2}	*-5.53*	*3.47*	*1.33*	*-3.46*
α_{-3}	3.44	8.18	3.84	12.72
α_{-4}	12.63	13.16	13.10	13.95
α_{-5}	19.45	14.06	15.29	23.68
α_{-6}	17.94	21.73	19.65	16.38
α_{-7}	-	31.92	26.69	-
β_{-1}	*-8.32*	*1.97*	*-5.71*	*0.15*
β_{-2}	34.95	13.38	22.54	18.08
$\alpha_1+\gamma_1$	12.20	8.97	11.17	7.39
$\alpha_2+\gamma_2$	37.51	13.74	18.88	36.17
$\alpha_3+\gamma_3$	-	24.35	21.48	-
VUNC (uncertainty)	2.25	*1.11*	2.69	*-1.66*

Monetary values by income and gender (ML)

Monetary values	Income		Gender	
	Low[a]	High	Male	Female
	ML	ML	ML	ML
α_{-1}	*-8.36*	*4.37*	*-4.74*	*0.61*
α_{-2}	-11.88	*0.35*	*-4.25*	-6.98
α_{-3}	*2.35*	8.22	*2.06*	9.64
α_{-4}	13.79	12.51	11.35	13.84
α_{-5}	15.23	13.98	14.41	21.23
α_{-6}	16.50	18.37	17.16	*10.42*
α_{-7}	-	32.76	33.43	-
β_{-1}	*-9.35*	*2.00*	*-6.75*	*0.32*
β_{-2}	14.61	*3.69*	8.02	*5.54*
$\alpha_1+\gamma_1$	6.18	-6.83	7.69	*3.01*
$\alpha_2+\gamma_2$	24.90	9.43	11.90	23.83
$\alpha_3+\gamma_3$	-	18.47	17.18	-
VUNC (uncertainty)	4.51	*1.00*	2.98	*2.64*

Note: [a]Low incomes are defined as yearly household gross income of less than €45,000.

CHAPTER 6

6 VALUATION OF TRAVEL TIME RELIABILITY FOR RAILWAY PASSENGERS

6.1 INTRODUCTION

Train reliability in terms of travel time is regarded as an important factor in train users' trip planning and decision making concerning their transport mode choice. The traveler's willingness-to-pay (WTP) for enhancing train reliability (i.e. the value of reliability (VOR)) will therefore play a large role not only in cost-benefit analyses of railway improvement projects, but also as a determinant of the travel demand for railway trips. The research topic of valuing travel time reliability has been receiving considerable attention recently (e.g. Small et al., 1999, Brownstone and Small, 2005; Hollander, 2006; Batley, 2007). Nevertheless, as discussed in Chapter 2, it remains unclear what the most appropriate modeling approach (utility specification) is to measure the monetary VOR. As we can observe in real life situations, there are several direct and indirect outcomes that railway passengers may experience because of unreliable services in rail transport. The direct consequences include missing the connections, arriving earlier or later than the desired arrival time, and increased waiting time. More indirect effects may concern stress or anxiety caused by the unreliability.

To investigate the effects of unreliability and to appraise the representative value for travel time reliability, we develop a stated choice experiment (SCE) for the estimation of behavioral models of train users when facing unreliability. It is clear that the services of road and public transport have different features. Bates et al. (2001) and Rietveld et al. (2001) pointed out various differences in reliability between road and public transport:

- Public transport usually follows the schedule posted in the timetable;
- For public transport, departing earlier than the timetable schedule is rare and arriving earlier than the timetable is also rare;
- The public transport service usually runs regularly according to its frequency(s).

Therefore, it is important to take these features into account when developing an SCE for assessing the value of reliability in public transport. The details of our experiment will be discussed below in Section 6.2.1.

The aims of this chapter are threefold. First, we will present the SCE that is designed specifically for valuing travel time reliability and schedule delay in the context of rail transport. Next, different modeling approaches, as discussed in Chapter 2 (Section 2.2.3): namely, the mean-variance model, Small’s trip scheduling model, or a combination of these two models, will be considered in estimating the values of time, reliability and/or schedule delay. And, finally, we will investigate the variations of reliability and schedule delay estimates over different groups of travelers.

The chapter is organized as follows. Section 6.2 describes the design of the SCE that is used to develop the travelers’ behavior models with respect to unreliability. Data collection is then addressed in Section 6.3. Section 6.4 discusses the results from analyzing the choice data that were collected in our experiment. Specifically, we will test the impacts of different modeling approaches on the empirical estimates, and then investigate various covariate effects and unobserved heterogeneity. Finally, Section 6.5 gives the conclusions and directions of future research.

6.2 SURVEY

To estimate the choice models for behavioral responses under the presence of train unreliability, we created an within-mode stated choice experiment that studies the effects of travel time, cost, and unreliability on the travelers’ decision making. We constructed the choice alternatives in the experiment as close as possible to the respondent’s actual choice environment, so that the tradeoff in each choice situation is more meaningful to the respondent. The choice alternatives are based on the respondent’s reported characteristics of his or her actual travel behavior (i.e. travel time, departure and arrival times, etc.). A respondent thus receives a tailored version of the SCE based on his or her current travel experience.

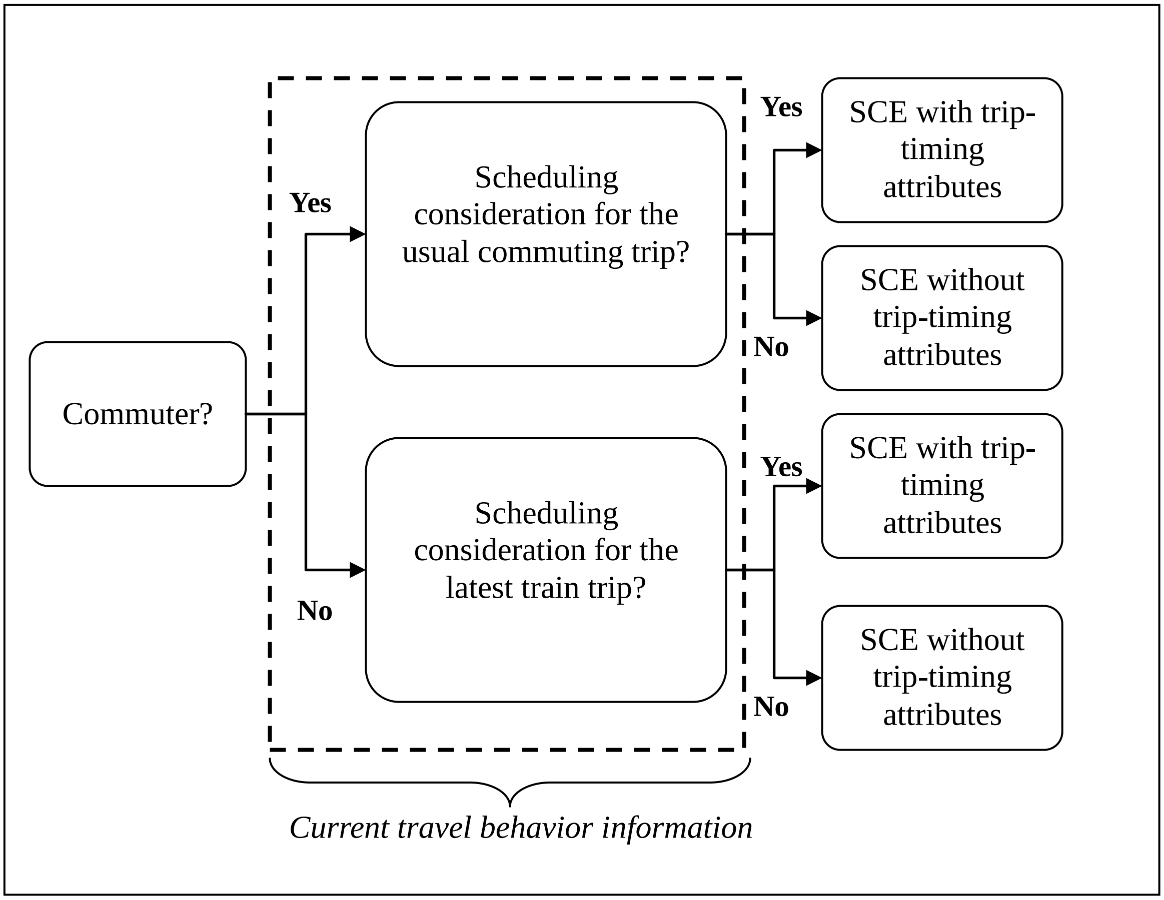


Figure 6.1 The structure of the survey with stated choice experiments (SCE)

The procedure of the survey is demonstrated in the flowchart shown in Figure 6.1. Basically, the questionnaire consists of two parts: the personal train trip experience and the individual customized within-mode stated choice experiment. We identify two types of train users, commuters and non-commuters, in the survey. The reason for doing so is that train commuters in general take the same or similar train connections (in the sense of the same (or nearby) origin and destination stations, departure and arrival times) in their daily commuting trips, while the non-commuters usually take different train connections to meet the needs or desires in some specific circumstances. From the literature, we note that a traveler's decision making for his or her scheduled trip may be very different from his or her non-scheduled trip. The main distinction between the scheduled and non-scheduled trips is the element of trip timing (departure and/or arrival times). For the scheduled trip, the departure and/or arrival times are planned in advance to meet the desired schedule of the traveler. So the timing of the trip is crucial for the scheduled trip, and the arrival and departure times desired by the traveler are called the preferred arrival and departure times. On the other hand, the non-scheduled trip refers to a trip where the timing is of less concern to the travelers. Leisure trips, such as for shopping, may be a good example. Because of the different nature of these two types of trips, we therefore develop two variants of SCEs to reflect the specific features of scheduled and non-scheduled train trips. The details of the SCE are discussed in the following subsection. Note that, for both work trips and non-work trips, we distinguish between scheduled and non-scheduled trips.

6.2.1 The stated choice experiment

For commuters or travelers who have scheduling considerations, the actual timing is more relevant than it is for non-scheduled trips. Instead of the service schedule (the train departure and arrival times), the service frequency may be more important to those train users who are less aware of the precise departure and arrival times in the timetable. By taking into account the different nature of these two types of trips, we set out different attributes for the two types of train users to make the SCE more sensible to the respondents. The scheduled trip SCE attributes are: departure time; timetable travel time; actual arrival time (implying unreliability); and fare; and the non-scheduled SCE attributes are: timetable travel time; train frequency; unreliability; and fare. In this section we discuss in turn each attribute of these two SCEs and explain how it is based on the information collected in the first part of the survey (see Fig. 6.1).

Timetable travel time

The timetable travel time is the travel time announced in the timetable, and attribute levels are calculated by multiplying the base travel time with a certain factor. The base (travel) time is the reported travel time of the respondent in the case of a reliable train connection. In the presence of train unreliability, the *timetable* travel time may not always be the *actual* travel time. Since the timetable travel time is posted publicly in the case of public transport, giving information about the timetable travel time rather than the actual or average travel time seems more intuitive in our experiment. Another advantage of showing the timetable travel time information is to make the respondents more aware of the differences between the travel time planned according to the timetable and the extra travel time due to the unreliable train services, so the resulting estimates from the choice responses can be interpreted more easily.

Train ticket price

The train ticket price is defined as the full price that a respondent has to pay for his one-way train journey. The fare level should ideally be based on a respondent's actual payment for his one-way journey. However, there exist some difficulties of cost computation for different types of tickets used by the respondents. For instant, some people use yearly public transport cards (paid by employers), the free student cards, or different kinds of discounted tickets. To overcome the above-mentioned difficulties, we related the respondent's self-reported travel time to the fare calculation in the experiment..Although the fare of public transport is usually set as a distance-based fee, in reality it is too risky to take the self-reported trip distance as the reference in this case. Travelers are usually more aware of the travel times than the distances, and this is particularly true for public transport users because travel times are posted in the timetable. In order to obtain more precise and credible information from the respondents, we therefore chose to use the travel time as the reference for the fare calculation, even though travel times and distances are not perfectly correlated in some cases[23].

Departure time

As discussed, trip timing is an important decision factor for travelers with scheduling considerations, it is therefore essential that the departure and arrival times serve as attributes in the scheduled trip experiment. In the first part of the questionnaire, we asked for individual's *preferred* departure time in the imaginary case that a reliable train service is guaranteed (without delay). When there is concern about unreliability, the traveler may either take an earlier train to reduce the possible consequences of delay or stick to the preferred departure time and then take the risk of being late. Therefore, we chose the self-reported preferred departure time as the latest ideal departure time, and varied the attribute levels by considering departures 15 or 30 minutes earlier than the preferred one.

Train service frequency

---

[23]For example when speeds of trains differ (e.g. stop train versus intercity train), or with different transfer times between train connections, travel times are certainly not perfectly correlated with travel distances.

A positive valuation of frequency, or reduced headway, for non-scheduled trips would suggest that travelers favor higher frequency of train services, even when that frequency itself does not help to achieve a higher reliability, or when it does not help to reduce scheduling costs. A higher frequency may just make travelers feel better: for example, when they consider flexibility or reliability aspects that are formally not varied between the alternatives in the experiment. It may thus be that the respondents implicitly interpreted the frequency as something that makes the trip better because it reduces the risk of trains being cancelled, or makes it possible to arrive at a more convenient time. So it seems to indicate that, even when travelers say that scheduling is not relevant, they still have some preference for being more flexible in the choice of departure time.

Arrival time / unreliability

One important feature of the reliability concept in public transport is adherence to the schedule (Bates et al., 2001). Although the discrepancy between the timetable and actual timing can be either in the departure or the arrival time, in the present study we will focus on the impact of unreliability on the arrival time only. Therefore, we assumed that the departure time of the train service is always reliable, and the consequence of an unreliable train connection affects the arrival time only. Meanwhile, by assuming that the train always departs reliably, we can also ignore the treatment of the possible prolonged waiting time at the departure station: all travel delays are incurred in-vehicle.

The presentation of travel time reliability is a key aspect of our experiment. As discussed before, reliability is referred to as the variation in possible travel times (or, a distribution of possible travel times), and can thus be expressed by a series of possible journey times/ arrival times/ delay times in the experiment. It is important that reliability has to be conveyed in a clear way for both dimensions: the events (journey times/arrival times/delay times), and the probabilities of the associated events. To simplify matters, and to ease the respondents' perception of the travel time variable, we used a two mass-points travel time distribution to depict the variability of travel time. In addition, to make the probability more understandable for the respondents, we used the 'x out of 10 trips' formulation to represent the probability of 10x percent (e.g. 1 out of 10 trips gives the probability of 10 percent). The formulation was modified several times according to the results and comments from the pilots of the survey. This is also consistent with the interview finding in

Chapter 3, in the sense that the verbal description is better understood than the graph by the respondents.

As mentioned, we assumed that the train always departs reliably, the arrival times and the delay times (unreliability) are therefore linearly dependent in our experiment. That is, the arrival time equals the departure time plus the sum of the timetable travel time and possible delay times, (unreliability). Given the information of timetable travel time and departure time in the choice profile, the arrival time and delay time can be inferred from each other. We decided to show the arrival time attribute to scheduled trip respondents to convey the information on unreliability, whereas we simply showed the possible delays to respondents without scheduling concerns.

Examples of the scheduled and non-scheduled trip choice sets, translated into English, are shown in Tables 6.1 and 6.2, respectively.

Table 6.1 Example of the choice question of the scheduled trip

Suppose your preferred arrival time at the destination station is ***<u>09.00 am</u>****, which of the 2 connections below would you prefer?*		
	☐ Connection A	**☐ Connection B**
Ticket price	8.6 EUR	5.9 EUR
Departure time	08:15 am	08:15 am
Timetable travel time	36 min	36 min
Arrival time	Always arrive at 08:51 am	08:51 am – 7 out of 10 trips 09:06 am – 3 out of 10 trips

Table 6.2 Example of the choice question of non-scheduled trip

Which of the 2 connections below would you prefer?		
	☐ Connection A	**☐ Connection B**
Ticket price	8.0 EUR	7.0 EUR
Timetable travel time	40 min	40 min
Train frequency	2 trains per hour	2 trains per hour
Unreliability	No delay	No delay: 8 out of 10 trips 10 min delay: 2 out of 10 trips

6.2.2 Design of the stated choice experiment

The heart of the stated choice experiment (SCE) is the experimental design (Louviere et al., 2000; Hensher et al., 2005). In the scheduled trips SCE, we have four attributes with 3, 3, 3, and 4 levels, respectively. It results in 108 (i.e. 3*3*3*4) possible treatment combinations[24] (full factorial) in the case of an unlabeled experiment design. Since the total of 108 treatment combinations is beyond what we can manage in the experiment, we select only part of the full possible combinations (fractional factorial design). We then pair the selected treatment combinations to construct the choice sets (choice tasks).

In pairing the treatment combinations, one practical strategy is to avoid the problem of dominance[25]. Because the nature of the SCE is to see respondents' tradeoffs between alternatives, it is important not to present the task where the preference is obvious. To facilitate the respondents' decision making, we kept one alternative (train connection) as always being reliable in a choice task. Furthermore, out of four attributes in one task, one or two attributes were held at the same levels between alternatives. Since the above-mentioned design strategies inevitably introduced some correlations between the attributes in the SCE[26], we therefore carried out the simulation exercise to make sure that the statistical design of the SCE is efficient enough to produce reliable estimates. In addition, we can check if the selected design and the level of attribute values can recover a sensible range of willingness-to-pay (WTP) for the attributes (i.e. the values of time (VOT), reliability (VOR), and schedule delay early/late (VSDE/VSDL)) through the simulation. The choice-set generating procedure was illustrated as follows:

1. Select the 'reliable' profiles (without delay) from all possible profiles;
2. Pair these selected 'reliable' profiles with 'unreliable' profiles;

---

[24] A treatment combination is the combination of attributes, each with unique levels; and is sometimes called a 'profile' in marketing research.
[25] The problem of dominance occurs when all the attributes in one alternative are all better than in the other. For instance, if train A has a lower price, shorter travel time, and more reliable service than train B, it is normal that train A would be chosen.
[26] In fact, retaining the orthogonality is not a guarantee of good design in our SCE. This is because the designed attributes of the SCE (e.g. departure time, reliability, etc.) are not the variables used in estimating the underlying utility function (e.g. schedule delay early and late). Thus, it is likely that zero correlations of the designed attributes can still produce high correlations among the variables used in the utility function.

3. Eliminate the alternative pairs where the preference is obvious (in the case of domination);
4. Select the design and decide the numbers of choice sets and blocks;
5. Develop values for the associated attribute levels;
6. Simulate responses and estimate models;
7. Review designs and values; if necessary, repeat 4.

In the simulation exercise (Step 6), we first generated some artificial choice responses based on the hypothetical coefficients (WTPs) and the statistical design (Steps 4 and 5). Next, we estimated a choice model from the artificial data, and compared the resulting coefficients and WTPs with the hypothetical ones. If the estimated coefficients and WTPs were not acceptable within a certain confidence level, we then went back to Step 4 to re-generate the choice tasks and/or to Step 5 to adjust the values/levels of the attributes. The procedure of Steps 4-6 was repeated until we obtained satisfactory estimates[27] for a sensible range of WTPs. Several designs produced satisfactory estimates, and we finally selected the one which yielded the smallest bias between the estimated and hypothetical coefficients and WTPs. Table 6.3 gives the resulting levels and values of the designed attributes.

Table 6.3 Overview of designed attributes and levels for the within-mode SCE

Unlabeled alternative	**Attributes**	**Levels**
Scheduled SCE	Timetable travel time Train ticket price Departure time Unreliability	3 (0.8*BTT; BTT; 1.2*BTT)[1] 3 (0.13*BTT; 0.16*BTT; 0.19*BTT) 3 (PDT-30; PDT-15; PDT)[2] 4 (100% of no delay; 90% of no delay & 10% of 15 mins delay; 70% of no delay & 30% of 15 mins delay; 80% of no delay & 20% of 30 mins delay)
Non-scheduled SCE	Timetable travel time Train ticket price Train frequency Unreliability	3 (0.8*BTT; BTT; 1.2*BTT) 3 (0.14*BTT; 0.16*BTT; 0.18*BTT) 2 (1 train per hour; 2 trains per hour) 4 (100% of no delay; 80% of no delay & 20% of 10 mins delay;

[27] The estimates are satisfactory when the true (hypothetical) coefficients are within the 95% confidence interval of the estimated ones.

		60% of no delay & 40% of 10 mins delay; 70% of no delay & 30% of 20 mins delay)

Notes: [1.]BTT is the base travel time reported from the respondents; [2.]PDT is the preferred departure time given from the respondents as well.

6.3 DATA COLLECTION

The web version of the survey was programmed by the online questionnaire company NetQuestionnaire. Special attention was paid to the routing of different groups of respondents, and to the construction of the personalized SCE. There were three pilot tests before the formal execution. The first pilot was a paper-based version, which was used to examine the wordings of the designed questions, the comprehensibility of the SCE, and the capability of estimating models from the choice responses. The second and third pilots were online web versions, which were used to test the functioning of the programmed routine and the personalized SCE. Meanwhile, we also collected respondents' comments through some follow-up questions at the end of the pilots. Many useful and constructive comments were considered and incorporated in the formal survey.

For the administration of the questionnaire, we cooperated with the Dutch Railway data collection agency, by using their panel of train users in the Netherlands. The data were collected in the period January to April 2006. Of the 5000 URL-linked emails sent, there were 2018 responses, of which 1756 questionnaires were completed. The response rate was about 40.4 percent (or 35.1 percent for completed questionnaire), which is considered high in this type of email questionnaire. We will report the results of the choice analyses later in Section 6.5.

6.4 ESTIMATION RESULTS FROM THE CHOICE DATA

Having reviewed the modeling approaches in Section 2.2.3 of Chapter 2, we now apply these specifications to our choice data. The data were analyzed separately for the scheduled and non-scheduled choice experiments. We start the analyses by multinomial logit (MNL) models, which are the benchmark of the choice models, and then analyze the effects of the observed individual heterogeneity by including some covariates into the MNL models. Finally, we advance the analyses by estimating some mixed logit (ML) models, which allow us to deal with the feature of panel data structure in the SCE.

6.4.1 Model specifications

Model specification of scheduled trip experiment

As a starting point, we estimate a model for the overall tradeoffs between travel time reliability and two other important trip attributes: namely, the expected travel time and the monetary cost. Using the standard deviation of travel time as a measure of travel time reliability, the model illustrates the central idea of the 'mean-variance' approach (Jackson and Jucker, 1981), where both the mean and the variance (standard deviation) of travel time play roles in a traveler's decision making. The model is given in Eq.(6.1):

$$U = \beta_C * C + \beta_\alpha * E[T] + \beta_\sigma * STD, \tag{6.1}$$

where *C* is the monetary travel cost, *E[T]* is the expected travel time, and *STD* is the standard deviation of travel time, representing travel time reliability.

We then move on to estimate a more complete model incorporating the scheduling variables based on the specification proposed by Noland and Small (1995). This scheduling model, given in Eq.(6.2), illustrates that the travelers account for the following attributes in their decision making: monetary travel cost *C*; expected travel time *E[T]*; expected schedule delay early *E[SDE]*; and expected schedule delay late *E[SDL]*:

$$U = \beta_C * C + \beta_\alpha * E[T] + \beta_\beta * E[SDE] + \beta_\gamma * E[SDL]. \tag{6.2}$$

The model shown in Eq.(6.2) assumes that the scheduling consideration is the sole reason for disliking unreliability, as there is no other reliability-related variable added in this model. Although this hypothesis could be true in some cases, it is more likely that people have some other aversions to unreliability, such as inconvenience, stress, or anxiety. To account for these mental effects of unreliability, we added the variable of standard deviation of travel time to Eq.(6.2), the scheduling model, so that we can examine if there is any additional disutility of unreliability apart from scheduling considerations. Eq.(6.3) shows the model specification and is called the ‘Full’ model, since it is a combination of the mean-variance and the scheduling model:

$$U = \beta_C * C + \beta_\alpha * E[T] + \beta_\beta * E[SDE] + \beta_\gamma * E[SDL] + \beta_\sigma * STD. \quad (6.3)$$

Model specification of non-scheduled trip experiment

The idea of model specification of the non-scheduled trip experiment is similar to the one in the scheduled trip experiment. The only difference is that scheduling variables are now replaced by a headway variable, which captures the preference of the non-scheduled trip travelers for a shorter headway (or higher frequency). In Eq.(6.4), the systematic indirect utility is given as:

$$U = \beta_C * C + \beta_\alpha * E[T] + \beta_\sigma * STD + \beta_H * H, \quad (6.4)$$

where H represents the train headway, or the inverse of train frequency.

6.4.2 Results of the multinomial logit (MNL) models

The models of the scheduled trips, based on Eq.(6.1)-Eq.(6.3), are estimated separately for the subgroups of commuters and non-commuters. For the choice data of non-scheduled trips, we pool the observations of the commuters and non-commuters, since there are too few commuters (only 27 respondents) in this subgroup. The MNL estimates for Eq.(6.1)-(6.4) are shown in Table 6.4, and the computed value of time (VOT), value of schedule delay early (VSDE), value of schedule delay late (VSDL), and value of reliability (VOR) are summarized in Table 6.5. Here we also derive the ratio between the coefficients of reliability and travel time. This ratio is called the ‘reliability ratio’ (RR), first defined by Black and Towriss (1993), which probably is a reasonable tool for comparison with prior studies. The resulting RRs from these models are listed in the last row of Table 6.5.

The overall estimates are plausible given their expected sign and reasonable values. In the case of the scheduled trips SCE, there is a substantial improvement of model fit by accounting for the scheduling costs. The result of LR test suggests that the scheduling full model outperforms in the model fit than the other models do at any statistical level of significance. Hence, it appears that the scheduling cost approach is the correct one in modeling this type of trip. Furthermore, the negative and significant coefficient of the reliability variable indicates that there are sources of disutility associated with unreliability other than pure scheduling considerations. Another reason why we prefer scheduling full models to mean-variance models is the plausible time values. The VOT estimates from the scheduling and full models are more comparable with the previous Dutch studies. According to AVV[28], the VOT appraisal used for the commuting train trip in 2006 is around €8.63 per hour, which is considerably lower than the estimates from the mean-variance model. It seems that, here also, ignoring the scheduling costs in the utility function may lead to biased (and higher) estimates of the time variable in the case of scheduled trips.

When comparing the estimates between commuters and non-commuters, we find that commuters' VOT, VSDE, and VSDL are generally higher than they are for non-commuters. This suggests that time and scheduling are more costly in commuting trips, than in trips for other purposes. These results are in line with what is generally found in the literature. Another plausible point to note is the relative sizes of the VOT, VSDE, and VSDL obtained from the scheduling and full models. Our result is in line with previous studies where the relationship VSDE < VOT < VSDL was found in most cases (see also Chapter 2). More in particular, the VSDL is more than double the VOT, and the VSDE is less than half the VOT.

It may also be interesting to compare the estimated results with the meta-analyses in Chapter 2. In the case of commuters' scheduled trips, the resulting RR in the full model is very close to the implied RR given by the same utility specification in Table 2.9 (0.36 versus 0.35). This indicates that the estimated result of the Full model for commuters' scheduled trips is in line with what we have found in the literature.

In the case of the non-scheduled trips SCE, what surprises us is that the resulting VOT (5.26 €/hour) is exceptionally close to the Dutch official VOT appraisal used for other purposes of train trips (according to AVV, this value is around 5.32 €/hour in 2006). This VOT is also comparable with the VOT derived from the non-commuters' scheduled trip SCE. The explanation for this is that most of the observations of the non-scheduled trip SCE are of non-commuters in our survey.

[28] Obtained from: http://www.rijkswaterstaat.nl/dvs/Images/personen%20vervoer%20trein_tcm178-167644.pdf.

As one might expect, the value of headway (VOH) is less than the VOT, and is estimated to be around one-third to one-quarter of the VOT in our data.

Table 6.4 MNL estimation results of the choice data of scheduled and non-scheduled SCE

Explanatory variables	**Scheduled trips**						**Non-scheduled trips**
	Mean-variance model		Scheduling model		Full model		
	Commuters	Non-commuters	Commuters	Non-commuters	Commuters	Non-commuters	Commuters and non-commuters
Travel cost C	-0.3343* (-17.374)	-0.2996* (-17.506)	-0.3719* (-18.713)	-0.2782* (-17.911)	-0.4521* (-19.409)	-0.3183* (-17.656)	-0.3530* (-13.838)
Mean travel time E[T] E[T]	-0.0870* (-24.040)	-0.0519* (-16.815)	-0.0765* (-16.732)	-0.0241* (-5.381)	-0.0883* (-18.152)	-0.0365* (-7.003)	-0.0310* (-15.282)
Expected schedule delay early E[SDE]			-0.0274* (-8.295)	0.0085* (2.001)	-0.0353* (-10.128)	-0.0007 (-0.150)	
Expected schedule delay late E[SDL]			-0.2218* (-24.811)	-0.0813* (-14.707)	-0.2150* (-24.713)	-0.0776* (-13.531)	
Reliability (standard deviation) R	-0.0514* (-13.178)	-0.0574* (-9.492)			-0.0321* (-6.762)	-0.0342* (-4.826)	-0.0685* (-8.658)
Train headway H							-0.0089* (-3.817)
N respondents Log likelihood	887 -4484.28	399 -1917.98	887 -4027.39	399 -1821.21	887 -4005.34	399 -1809.42	446 -2297.85

Note: t-statistics are shown in parentheses. Significance is indicated by *, referring to the 95% significance level.

Table 6.5 Monetary values and reliability ratios implied by Table 6.4 (€/hour)

	Scheduled trips						**Non-scheduled trips**
	Mean-variance model		**Scheduling model**		**Full model**		
	Commuters	Non-commuters	Commuters	Non-commuters	Commuters	Non-commuters	Commuters and non-commuters
VOT	**16.62**	**10.40**	**12.35**	**5.2**	**11.71**	**6.88**	**5.26**
VSDE			**4.43**	**-1.84**	**4.68**	0.13	
VSDL			**35.78**	**17.54**	**28.53**	**14.63**	
VOR	**9.22**	**11.50**			**4.26**	**6.44**	**11.64**
VOH							**1.51**
RR	**0.59**	**1.11**			**0.36**	**0.94**	**2.21**

Note: The estimates with a 95% significance level are printed in bold print.

It can be seen that the reliability ratios in full models are smaller than in the mean-variance models. The reason is intuitive, since part of the disutilities of unreliability is now picked up by the scheduling costs in the case of scheduled trips. It is striking to see that the RR in the non-scheduled trips is much higher than it is in the scheduled ones, although the VORs in these two cases (see the VOR in the mean-variance model for scheduled trips) are comparable with each other. It is not easy to think of a solid explanation for this finding.

6.4.3 Results of observed heterogeneity in multinomial logit models: covariate effects

The prior literature has shown that the VOTs and schedule delay vary with some trip characteristics and socioeconomic variables, such as income and gender (e.g. Small et al., 1999; Lam and Small, 2001, Chapter 2). In this subsection we investigate the effects of income, gender, education, and travel cost compensation by the employer on our estimates of interest. By specifying interaction terms of the attributes with some socioeconomic dummies in the utility function, we are able to identify whether there are statistically significant differences between these groups. Eq.(6.5) gives an example of the model specification for investigating the impact of income:

$$
\begin{aligned}
U = {} & \beta_C * C + \beta_T * E[T] + \beta_{SDE} * E[SDE] + \beta_{SDL} * E[SDL] + \beta_R * STD \\
& + \beta_{C\_Hinc} * Hinc * C + \beta_{T\_Hinc} * Hinc * E[T] + \beta_{SDE\_Hinc} * Hinc * E[SDE] \\
& + \beta_{SDL\_Hinc} * Hinc * E[SDL] + \beta_{R\_Hinc} * Hinc * STD,
\end{aligned} \tag{6.5}
$$

where $Hinc$ =1 if the monthly household income is above €3,250, and 0 otherwise.

Thus, by checking the significance of the coefficients of these interaction terms, we can examine whether the estimates are significantly different between low and high income groups. The estimation results can be found in Appendices 6A, 6B, and 6C, and the implied monetary values are summarized in Table 6.6.

Table 6.6 Implied monetary values from different sub-groups (€/hour)

Scheduled trips for commuters												
	Income[a]		**Education[b]**		**Gender**		**Trip lengths[c]**			**Travel cost compensation**		
	Low	High	Low	High	Male	Female	Short	Medium	Long	No	Partial	Fully
VOT	11.63	12.00	**8.26**	**12.59**	**12.11**	**10.93**	**7.03**	**11.90**	**17.20**	**4.02**	**11.73**	**14.21**
VSDE	4.47	4.94	4.45	4.61	**4.49**	**4.82**	**3.44**	**4.60**	**5.31**	**3.23**	**4.93**	**5.78**
VSDL	27.56	32.45	29.82	28.26	**29.75**	**27.50**	**21.30**	**24.36**	**28.98**	**21.43**	**25.34**	**38.4**
VOR	3.10	5.77	3.41	4.37	**5.07**	**3.61**	**2.12**	**4.24**	**9.01**	**4.31**	**6.59**	**7.72**
Scheduled trips for non-commuters												
	Income		**Education**		**Gender**		**Trip lengths**					
	Low	High	Low	High	Male	Female	Short	Medium	Long			
VOT	**5.93**	**10.9**	**4.49**	**8.98**	**8.4**	**5.47**	(0.54)	(7.58)	(7.79)			
VSDE	(-0.97)	2.76	(-1.51)	(0.83)	(0.06)	(-0.09)	**1.53**	2.01	**0.21**			
VSDL	**13.59**	**23.57**	14.85	15.20	**17.24**	**11.22**	**13.13**	**12.45**	**15.12**			
VOR	5.07	11.21	**(2.88)**	**6.24**	**9.8**	**6.38**	**5.23**	**4.03**	**5.30**			
Non-scheduling trips												
	Income		**Education**		**Gender**		**Trip lengths**					
	Low	High	Low	High	Male	Female	Short	Medium	Long			
VOT	**4.85**	**6.92**	**4.30**	**6.83**	5.90	4.80	**1.70**	**4.92**	**5.54**			
VOH	1.70	0.79	**(0.83)**	**1.64**	1.43	1.61	**0.91**	**1.11**	**1.27**			
VOR	11.30	13.28	12.54	10.72	**15.92**	**9.59**	**3.27**	**9.29**	**12.15**			

Notes: Bold print is used where covariates are significant at the 95% significance level. The values shown in parentheses are the estimates, which are not significantly different from zero at the 95% significance level.

[a] The high income group refers to those whose household monthly net income are above €3,250.

[b] The high education group refers to those whose education level are HBO or above.

[c] The short, medium, and long trip lengths refer to 0-40km, 40-80km, and ≥ 80km, respectively.

The results of these covariate effects are plausible. As shown in Table 6.6, the variations of the estimates are similar for the three types of trips (scheduled commuting, scheduled non-commuting, and non-scheduled trips). In general, income has a positive effect on the VOT and the VOR, though sometimes the effect is insignificant. Education, as expected, has a similar effect to income. Female railway users have a lower VOT and VOR in our data, which is different from some earlier findings (Lam and Small, 2001). The effect of trip length has a significant positive impact on the VOT and the VOR, in particular for the groups of scheduled commuting trips and non-scheduled trips. For scheduled commuting trips, the VOT, VSDE, VSDL, and VOR increase considerably with the levels of cost compensation that commuters can obtain from their employers. This is probably the consequence of people anticipating compensation at higher prices when already receiving compensation, which would reduce the cost coefficient in the estimate. Travel cost compensation thus has a great impact on the valuation of the travel time, schedule delay and reliability estimates, and this impact is even stronger than other effects such as income, education, and gender for commuters.

6.4.4 Results of mixed logit (ML) models

A number of different mixed logit (ML) model specifications were estimated in our analysis. The utility function we used is based on Eq.(6.3) for scheduled trips and Eq.(6.4) for non-scheduled trips. Given the presence of multiple observations that were drawn from the same individual (which is a typical feature in SCEs), the models we estimated here account for the correlated responses given by each sampled individual (Revelt and Train, 1998).

Similar to the MNL estimations in Section 6.5.2, we estimated separate models for commuter scheduled trips, non-commuter scheduled trips, and non-scheduled trips. The final preferred ML models take the form of one random parameter on the reliability variable[29]. For the samples of commuters' scheduled trips, since we found the presence of the cost compensation and the trip length measures (observed individual heterogeneity) has a strong influence on the cost and travel time parameters, we therefore selected the model with these observed covariates as our final

[29] With the present data set, the mixed logit model hardly converged if we selected other variables or added more variables as the random parameters in the estimation.

preferred model[30]. We also tried different types of distribution (e.g. normal, uniform, triangular, and log-normal) for the random parameter. Except for the case of log-normal, which hardly converged, the rest of the distributions result in rather similar estimates in the models. Thus, owing to a slightly better fit measure in the estimation, the models with a normally distributed random parameter were chosen to be presented here. The results of the ML models for these three types of trips are summarized in Table 6.7.

The implied monetary values for non-commuter scheduled trips and non-scheduled trips are given in the bottom rows of Table 6.7, and these values for commuter scheduled trips are given in Table 6.8 separately. The significant and large standard deviation of the random parameter in the ML estimations suggests that there is great level of unobserved heterogeneity among individuals with respect to the valuation of travel time reliability. The monetary values obtained from MNL models and the mean values derived from ML models are quite comparable. The resulting monetary values of different trip lengths and cost compensation levels for scheduled trips commuters from the ML models are also similar to the variations pattern observed in the MNL models with covariates. In general, the commuters have higher values of time and reliability as their commuting distances increase and/or the levels of cost compensation rise. These findings suggest that the overall analyses in both the MNL and ML models are consistent and the derived monetary values are robust for different types of travelers and trips.

[30] The other reason why we chose the model with observed covariates as the final preferred one is that the model without covariates did not lead to convergence in our data set.

Table 6.7 ML estimation results of the choice data

Explanatory variables	**Scheduled trips for commuters** Coeff.	t-stat.	**Scheduled trips for non-commuters** Coeff.	t-stat.	**Non-scheduled trips** Coeff.	t-stat.
Random parameter mean effects (normal distribution)						
Reliability, R	-0.0590*	-8.04	-0.0474*	-3.95	-0.0941*	-6.28
Random parameter standard deviation						
Reliability, R	0.1276*	16.78	0.1574*	13.04	0.2261*	15.54
Observed heterogeneity						
Travel cost, C: Partial cost compensation	-0.1044*	-2.25				
Travel cost, C: Full cost compensation	-0.2260*	-2.26				
Travel cost, C: trip distance between 40-80km	0.4642*	6.49				
Travel cost, C: trip distance ≥ 80km	0.7563*	10.16				
Mean travel time, E[T]: Partial cost compensation	-0.0095	-1.08				
Mean travel time, E[T]: Full cost compensation	0.0515*	3.74				
Mean travel time, E[T]: trip distance between 40-80km	0.0249*	2.04				
Mean travel time, E[T]: trip distance ≥ 80km	0.0424*	3.27				
Non-random parameters						
Travel cost, C	-1.0074*	-14.13	-0.3717*	-16.79	-0.4796*	-14.26
Mean travel time, E[T]	-0.1385*	-11.59	-0.0428*	-7.27	-0.0410*	-15.46
Expected schedule delay early, E[SDE]	-0.0444*	-11.61	-0.0022	-0.41		
Expected schedule delay late, E[SDL]	-0.2568*	-24.69	-0.0947*	-14.24		
Reliability, R						
Train headway, H					-0.0102*	-3.34
N respondents	887		399		466	
Log-likelihood	-3817.99		-1716.16		-2473.15	
Pseudo-R-sqrd	0.2223		0.2229		0.1359	
Mean Monetary values (€/ hour)						
VOT	(see Table 6.8)		**6.91***		**5.13***	
VSDE	(see Table 6.8)		0.35			
VSDL	(see Table 6.8)		**15.28***			
VOR	(see Table 6.8)		**7.65***		**11.77***	
VOH					**1.27***	

Note: Significance is indicated by *, referring to the 95% significance level.

Table 6.8 Monetary values of different trip lengths and cost compensation levels for scheduled trips for commuters in Table 6.7 (€/hour)

	No-cost compensation			**Partial-cost compensation**			**Full-cost compensation**		
	0-40km	**40-80km**	**≥ 80km**	**0-40km**	**40-80km**	**≥ 80km**	**0-40km**	**40-80km**	**≥ 80km**
VOT	4.23	4.84	5.60	7.99	11.40	17.83	8.25	12.54	22.96
VSDE	2.16	3.47	5.59	2.40	4.12	7.50	2.65	4.92	10.62
VSDL	12.49	20.03	32.29	13.86	23.79	43.34	15.29	28.36	10.62
VOR	2.87	4.61	7.43	3.19	5.47	9.97	3.52	6.52	14.11

6.5 CONCLUSIONS

A number of main conclusions can be drawn from the foregoing analyses. First, the overall parameters estimated from the choice models are plausible, and so are the implied values of time and reliability. The VOT estimates obtained from our data set are comparable to the current official Dutch appraisals. The results from the various models suggest that people do care a lot about the reliability of train services, no matter what types of trips they are making, and in some cases reliability is evaluated higher than the travel time. The second conclusion is that the monetary values of time, schedule delay, and reliability vary considerably over different subgroups of respondents. The results of the covariate analyses are rather plausible. Most of the findings are in line with the literature, and the variations of the estimates are also consistent for three subsets of our data. Third, the estimation results of the ML models indicate that there is a strong degree of unobserved heterogeneity of the reliability of travel time among individuals. As the monetary values derived from the ML models are comparable to the values obtained from MNL models, this suggests that our overall analyses and modeling framework are quite robust.

One important finding in this chapter is that imposing the mean-variance model may lead to biased estimates of VOT and VOR in the scheduled trips. A recent UK VOR study of bus users (Hollander, 2006) concluded that the scheduling model is preferred to the mean-variance model in terms of modeling the effect of travel time reliability. Here, we recommend using the model in which both scheduling and reliability variables are fully specified in the utility function, because such a model gives better explanatory power with respect to travelers' choice behavior for this type of trip. However, it should be noted that the resulting monetary values of schedule delay are difficult to implement in a cost-benefit analysis of a transportation project, given that the preferred arrival times are usually unknown by the policy makers. Thus, the questions how to make a link between the schedule delay early/late and the reliability (standard deviation of travel time), and how to convert the cost of scheduling to a generalized cost of reliability become relevant in practical application. In the next chapter, we will discuss these issues in more detail.

Appendix 6A: MNL results with covariate effects: Scheduled trips for commuters

Explanatory variables	Income		Education		Gender		Trip lengths		Travel cost compensation	
	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.	Est. B	t-stats.	Est. b	t-stats.
Travel cost, C	-0.4531*	-15.133	-0.4546*	-9.145	-0.4064*	-14.890	-0.9987*	-15.012	-0.7033*	-7.685
C*High_income (≥ 3,250 monthly)	0.0270	0.577								
C*High_education (HBO or above)			0.0105	0.188						
C*Female					-0.1637*	-3.358				
C*medium trip length (40-80 km)							0.4991*	6.680		
C*long trip length (≥ 80 km)							0.7634*	10.091		
C*Partial compensation									0.3108*	3.199
C*Fully compensation									0.2438*	2.485
Mean travel time, E[T]	-0.0879*	-13.903	-0.0626*	-6.739	-0.0820*	-14.109	-0.1170*	-10.887	-0.0471*	-3.293
E[T]*High_income (≥ 3,250 monthly)	0.0026	0.269								
E[T]*High_education (HBO or above)			-0.0327*	-3.006						
E[T]*Female					-0.0218*	-2.159				
E[T]*medium trip length (40-80 km)							0.0179	1.381		
E[T]*long trip length (≥ 80 km)							0.0495*	3.139		
E[T]*Partial compensation									-0.0458*	-2.885
E[T]*Full compensation									-0.0427*	-2.628
Expected schedule delay early, E[SDE]	-0.0338*	-7.490	-0.0337*	-4.986	-0.0304*	-7.097	-0.0573*	-10.606	-0.0378*	-3.301
E[SDE]*High_income (≥ 3,250 monthly)	-0.0013	-0.187								
E[SDE]*High_education (HBO or above)			-0.0012	-0.153						
E[SDE]*Female					-0.0154*	-2.150				
E[SDE]*medium trip length (40-80 km)							0.0191*	2.444		
E[SDE]*long trip length (≥ 80 km)							0.0365*	3.133		
E[SDE]*Partial compensation									-0.0043	-0.345
E[SDE]*Full compensation									0.0129	1.021
Expected schedule delay late, E[SDL]	-0.2081*	-18.698	-0.2259*	-12.923	-0.2015*	-19.294	-0.3546*	-19.727	-0.2512*	-8.613
E[SDL]*High_income (≥ 3,250 monthly)	-0.0223	-1.237								
E[SDL]*High_education (HBO or above)			0.0118	0.587						
E[SDL]*Female					-0.0598*	-3.284				
E[SDL]*medium trip length (40-80 km)							0.1517*	7.088		
E[SDL]*long trip length (≥ 80 km)							0.2409*	10.667		
E[SDL]*Partial compensation									0.0174	0.545
E[SDL]*Full compensation									0.0571	1.789
Reliability (standard deviation), R	-0.0234*	-3.749	-0.0258*	-2.768	-0.0343*	-5.626	-0.0353*	-4.714	-0.0505*	-3.052
R*High_income (≥ 3,250 monthly)	-0.0175	-1.739								
R*High_education (HBO or above)			-0.0073	-0.664						
R*Female					0.0087	0.850				
R*medium trip length (40-80 km)							-0.0055	-0.500		
R*long trip length (≥ 80 km)							-0.0168	-0.935		
R*Partial compensation									0.0162	0.906
R*Full compensation									0.0301	1.646
Log-likelihood	-4003.32		-4000.75		-4004.25		-3903.19		-3981.35	

Note: t-statistics are shown in parentheses. Significance is indicated by *, referring to the 95% significance level.

Appendix 6B: MNL results with covariate effects: Scheduled trips for non-commuters

Explanatory variables	Income		Education		Gender		Trip lengths	
	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.
Travel cost, C	-0.3031*	-15.252	-0.2957*	-10.781	-0.2493*	-10.368	-1.3104*	-7.860
C*High_income (≥ 3,250 monthly)	-0.0598	-1.325						
C*High_education (HBO or above)			-0.0293	-0.815				
C*Female					-0.1338*	-3.726		
C*medium trip length (40-80 km)							0.8096*	4.600
C*long trip length (≥ 80 km)							1.0487*	6.250
Mean travel time, E[T]	-0.0300*	-5.152	-0.0221*	-2.803	-0.0349*	-4.619	-0.0117	-0.462
E[T]*High_income (≥ 3,250 monthly)	-0.0251	-1.990						
E[T]*High_education (HBO or above)			-0.0221*	-2.124				
E[T]*Female					-0.0006	-0.059		
E[T]*medium trip length (40-80 km)							-0.0516	-1.844
E[T]*long trip length (≥ 80 km)							-0.0224	-0.844
Expected schedule delay early, E[SDE]	0.0049	0.929	0.0075	1.004	-0.0002	-0.034	-0.0335*	-2.929
E[SDE]*High_income (≥ 3,250 monthly)	-0.0216	-1.951						
E[SDE]*High_education (HBO or above)			-0.0119	-1.252				
E[SDE]*Female					0.0008	0.085		
E[SDE]*medium trip length (40-80 km)							0.0167	1.111
E[SDE]*long trip length (≥ 80 km)							0.0326*	2.428
Expected schedule delay late, E[SDL]	-0.0686*	-10.901	-0.0932*	-8.617	-0.0716*	-8.776	-0.2869*	-7.637
E[SDL]*High_income (≥ 3,250 monthly)	-0.0504*	-3.258						
E[SDL]*High_education (HBO or above)			-0.0092	-0.797				
E[SDL]*Female					-0.0144	-1.237		
E[SDL]*medium trip length (40-80 km)							0.1830*	4.497
E[SDL]*long trip length (≥ 80 km)							0.2210*	5.804
Reliability (standard deviation), R	-0.0256*	-3.023	-0.0142	-1.235	-0.0407*	-3.716	-0.1143*	-6.560
R*High_income (≥ 3,250 monthly)	-0.0310	-1.838						
R*High_education (HBO or above)			-0.0307*	-2.066				
R*Female					0.0129	0.877		
R*medium trip length (40-80 km)							0.0807	3.581
R*long trip length (≥ 80 km)							0.0912	4.370
Log likelihood	-1801.20		-1806.40		-1798.95		-1753.86	

Note: t-statistics are shown in parentheses. Significance is indicated by *, referring to the 95% significance level.

Appendix 6C: MNL results with covariate effects: Non-scheduled trips

Explanatory variables	Income		Education		Gender		Trip lengths	
	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.	Est. b	t-stats.
Travel cost, C	-0.3615*	-12.602	-0.3560*	-10.069	-0.3383*	-9.509	-2.3196*	-7.321
C*High_income (≥ €3,250 monthly)	0.0330	0.513						
C*High_education (HBO or above)			-0.0109	-0.212				
C*Female					-0.0454	-0.882		
C*medium trip length (40-80 km)							1.5163*	4.560
C*long trip length (≥ 80 km)							2.0087*	6.313
Mean travel time, E[T]	-0.0292*	-13.444	-0.0255*	-9.982	-0.0333*	-11.370	-0.0658*	-3.027
E[T]*High_income (≥ €3,250 monthly)	-0.0125*	-2.071						
E[T]*High_education (HBO or above)			-0.0150*	-3.533				
E[T]*Female					0.0035	0.857		
E[T]*medium trip length (40-80 km)							-0.0028	-0.122
E[T]*long trip length (≥ 80 km)							0.0371	1.698
Headway, H	-0.0102*	-3.890	-0.0049	-1.542	-0.0081*	-2.321	-0.0351*	-5.565
H*High_income (≥ € 3,250 monthly)	0.0059	1.007						
H*High_education (HBO or above)			-0.0097*	-2.054				
H*Female					-0.0022	-0.473		
H*medium trip length (40-80 km)							0.0203*	2.555
H*long trip length (≥ 80 km)							0.0286*	4.028
Reliability (standard deviation), R	-0.0681*	-7.556	-0.0744*	-6.681	-0.0897*	-7.385	-0.1265*	-6.347
R*High_income (≥ € 3,250 monthly)	-0.0046	-0.241						
R*High_education (HBO or above)			0.0089	0.558				
R*Female					0.0357*	2.222		
R*medium trip length (40-80 km)							0.0021	0.079
R*long trip length (≥ 80 km)							0.0636*	2.723
Log likelihood	-2293.14		-2285.56		-2290.43		-2252.38	

Note: t-statistics are shown in parentheses. Significance is indicated by *, referring to the 95% significance level.

CHAPTER 7

7 THE ANALYSIS OF ANTICIPATING DEPARTURE AND THE VALUE OF IMPLIED SCHEDULE DELAY FOR RAILWAY PASSENGERS

7.1 INTRODUCTION

Departure time choice is an important element in travelers' decision making, and it becomes more complicated when the travel time is unreliable. As the degree of travel time unreliability increases, travelers can be expected to shift their departure times to earlier hours to compensate for the increased probability of being late (Gaver, 1968; Knight, 1974; Polak, 1987). How early a traveler would shift his departure schedule depends on how he makes tradeoffs among early arrivals, late arrivals, and travel time (in the case when travel time varies by time of day), for uncertain travel times. An optimal departure time can be expected to be chosen, such that the traveler's resulting expected utility is maximized (Noland and Small 1995).

In the context of road transport, Noland and Small (1995) developed an analytical model for car commuters' departure time choice that takes account of the effect of uncertain travel time. In this model, the optimal 'head-start' time and optimal expected cost function were derived under uniform and exponential travel time distributions, representing the probability distribution of time-varying travel times under different levels of congestion.

The choice of departure time is also of great interest in public transport. Unlike road transport, where the departure time can be chosen freely, the public transport traveler can only choose between some fixed scheduled services, defined by the timetable. Thus, it is common that train travelers may choose to take an earlier than necessary connection to reduce the probability of arriving too late. It is similar to the situation in road transport where people may reserve a 'safety margin' for unexpected delays in traffic, but in public transport the departure time is usually discrete and restricted by the timetable. In the present paper, we refer to the behavior of taking an earlier than strictly necessary train connection as "anticipating departure". We will develop a simple model to investigate how the traveler makes the decision between the scheduled services when unreliability is involved.

Since the reliability of travel time is becoming increasingly important in policy evaluation, it is important that the valuation of reliability can be incorporated into cost-benefit analyses (CBAs) of infrastructure projects and improvements. Nevertheless, it is not straightforward how monetary values obtained from schedule delay early and late can be used in a CBA framework in this context. Travel time period choices are often not incorporated in CBAs, and preferred arrival times are typically unknown (Hamer et al., 2005). Given that reliability is often valued in terms of the standard deviation of travel times[31], it would be useful if schedule delay cost parameters could be translated into an implied value, which is based on the standard deviation definition.

Bates et al. (2001) concluded from the model established by Noland and Small that the expected schedule delay cost can be well approximated by a functional form, which is linear to the standard deviation of travel time distribution in the case of car journeys. In addition, Bates et al. also considered some complex modeling issues for scheduled public transport services, where travelers can only choose from a set of head starts with a fixed interval. Fosgerau and Karlström (2007) attempted to derive travelers' expected cost function in relation to the standard deviation of travel time distribution. They obtained a positive result that the optimal expected cost is linear in the mean and standard deviation of travel time distribution in the context of car journeys. Nevertheless, they were not able to find a general explicit solution for general travel time distributions in the case of scheduled services, but concluded that travelers' expected cost function is not linear in the mean and standard deviation of travel time distribution.

In contrast to the approach used in earlier studies, which often consider general distributions, we choose to derive the expected schedule delay cost for scheduled train services under some specific situations. In our framework, travelers' "anticipating departures" (departing earlier than strictly necessary in response to travel time unreliability) will also be taken into account. In such a case, the marginal expected scheduling costs associated with the standard deviation, defined as

[31] This was also proposed for the Netherlands at a workshop organized by RAND Europe and AVV (Oct. 2004).

the 'value of implied schedule delay' can also be obtained in the present study, under these hypothesized situations.

The outline of this chapter is as follows. Section 7.2 presents the theoretical model for departure time choice with scheduled services in public transport. Section 7.3 discusses anticipating behavior, and derives the value of implied schedule delay from a certain travel time distribution for railway passengers. Section 7.4 provides numerical illustrations based on the analytical framework developed in previous sections. In Section 7.5 we apply the results in the calculation of benefit contributions by incorporating the effects of reliability and/or schedule delay in a cost-benefit analysis. Section 7.6 then concludes.

7.2 THE MODEL

Suppose a train runs regularly with a headway H, and is scheduled to arrive at t=T. Assume that a traveler's preferred arrival time (PAT) is in the range [T, T+H], so he has to decide whether to catch the connection arriving at t=T, a later one arriving at t=T+H, or perhaps even another one. We assume that this traveler will choose the connection that gives him the maximum expected utility (or minimum expected disutility). In principal, if the train service is completely reliable, this traveler's choice would either be the service that arrives at T (denoted as train connection B) or the one at T+H (denoted as train connection C), depending on his PAT and the values of schedule delay early and late. Nevertheless, if the train service is very unreliable, (particularly when the possible delay time approaches the headway), it may be possible that this traveler chooses to take an earlier than necessary connection, with a timetable arrival time at t=T-H (denoted as connection A). Without loss of generality, we can set the clock time T=0. Figure 7.1 then illustrates travelers' PATs and the arrival times of train connections A, B, and C.

We want to study travelers' decisions under these circumstances. The analysis set up for this chapter makes the following assumptions:

1. The headway of the train service is H, and time T=0 is chosen such that the traveler's PATs are in the range [0, H];

2. The train scheduled to arrive at t=-H (timetable arrival time, not the actual arrival time) is called connection ‘A’; the train to arrive at t=0 is called connection ‘B’; and the train to arrive at t=H is called connection ‘C’;
3. Each train has a probability of being late. More precisely, there is a distribution of n possible delays DT_i (**D**elay **T**ime), each occurring with a probability p_i. We indicate a zero delay as $DT_{0.}$=0, and observe that $p_0 = 1 - \sum_{i=1}^{n} p_i$. We assume that $0 \le DT_i \le H$ [32].
4. Since the trip duration of the train service, unlike in road transport, is usually time-invariant, the travel time can also be normalized to zero without loss of generality. For a train without delay, the departure time and arrival time then coincide.

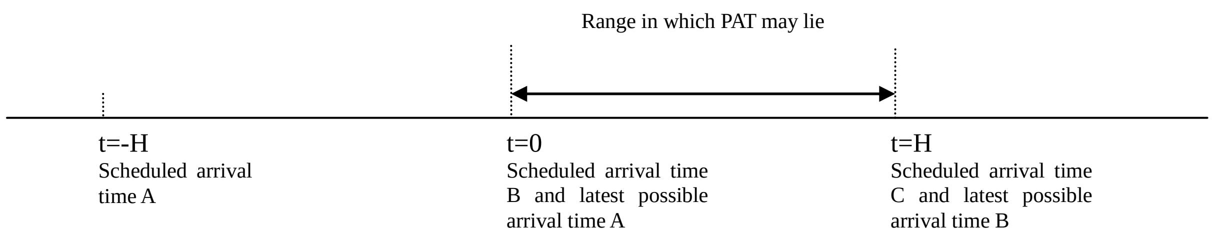


Figure 7.1 Travelers’ PATs and arrival times of different connections

To derive the relationship between H, PAT, and schedule delay parameters, we adopt the functional form proposed by Noland and Small (1995), which is extensively used in the literature of the departure time choice modeling (see, e.g., Ettema et al. 2005, Ettema and Timmermans, 2006), to express the scheduling disutility as:

$$DU_s = \beta \cdot E[SDE] + \gamma \cdot E[SDL] + \theta \cdot P_L, \qquad (7.1)$$

where DU_s denotes the disutility associated with scheduling; $E[SDE]$ is the expected schedule delay early; $E[SDL]$ is the expected schedule delay late; and P_L is the expected probability of late

---

[32] The assumption that $DT_i \le H$ facilitates the analysis of travelers’ anticipating departures. Since travelers may take the connection that arrives at t=-2H or earlier when the delay time DT is larger than 2H or more, abandoning this assumption would require consideration of a fourth possible service. Therefore, to simplify the analysis for predicting the anticipating departures, we consider the case for which the longest delay time is H, so that a traveler will only have to choose between connections A, B, and C.

arrival. The $E[SDE]$, $E[SDL]$, and P_L for train connections A, B, and C can be computed as follows:

Train connection A

$$\begin{aligned} E[SDE] &= \sum_{i=0}^{n} p_i \cdot \left[PAT + H - DT_i\right] \\ E[SDL] &= 0 \\ P_L &= 0. \end{aligned} \tag{7.2}$$

Train connection B

If $PAT \geq DT_n$, then

$$\begin{aligned} E[SDE] &= \sum_{i=0}^{n} p_i \cdot \left[PAT - DT_i\right] \\ E[SDL] &= 0 \\ P_L &= 0. \end{aligned} \tag{7.3}$$

Otherwise, define j such that $DT_{j-1} \leq PAT < DT_j$. Then:

$$\begin{aligned} E[SDE] &= \sum_{i=0}^{j-1} p_i \cdot \left[PAT - DT_i\right] \\ E[SDL] &= \sum_{i=j}^{n} p_i \cdot \left[DT_i - PAT\right] \\ P_L &= \sum_{i=j}^{n} p_i. \end{aligned} \tag{7.4}$$

Train connection C

$$\begin{aligned} E[SDE] &= 0 \\ E[SDL] &= \sum_{i=0}^{n} p_i \cdot \left[(H + DT_i) - PAT\right] \\ P_L &= 1. \end{aligned} \tag{7.5}$$

The decision rule is that a traveler will choose train connection k = [A, B, C] if the schedule disutility of train k, $DU_s(k)$, is the smallest of these three available train connections.

In the following section, we discuss how we apply this model to predict the anticipating departure behavior, and to calculate the expected scheduling cost and the value of implied scheduling costs given a certain distribution of travel times.

7.3 THE ANTICIPATING DEPARTURE BEHAVIOR AND THE IMPLIED SCHEDULING COSTS

7.3.1 Unreliable train service and travelers' anticipating departure

In public transport, it is common that people may choose an earlier than necessary connection, according to the timetable, in order to reduce the probability of late arrival. This kind of behavior is referred to as 'anticipating departure' in the present study. Here, we will explore the influence of unreliability on train users' anticipating departures, in the context of the conceptual framework presented in the previous section.

Let us first consider an individual's departure time choice in the case of a completely reliable train service. In that case, connection A will never be taken by a traveler, since it is always dominated by connection B (the schedule delay early (SDE) of connection A is always larger than it is for connection B). Therefore, a traveler will choose either connection B or C, depending on his PAT and schedule delay parameters. For the sake of simplicity, we assume that $\theta = 0$ here. The switching level of PAT, denoted PAT*, for a choice between two reliable train services can then be computed as follows:

$$DU_s\left(ConnectionB\right) = DU_s\left(ConnectionC\right)$$
$$\Rightarrow \beta \cdot (PAT^*) = \gamma \cdot (H - PAT^*)$$

$$\Rightarrow PAT^* = \frac{\gamma \cdot H}{\beta + \gamma}. \tag{7.6}$$

It implies that a traveler will choose to take connection B if his *PAT* falls in the interval $[0, \frac{\gamma \cdot H}{\beta + \gamma}]$, and will switch to connection C if his PAT is in the interval $[\frac{\gamma \cdot H}{\beta + \gamma}, H]$ under the reliable train service circumstance. This result and Equation (7.6) were also obtained by de Palma and Lindsey (2001) and Bates et al. (2001).

Based on the analysis above, we can also compute a traveler's connection choice behavior under travel time unreliability (given the amount of delay times and their associated probabilities). By comparing individuals' connection choices between unreliable services with the choices between reliable services, we are able to determine whether a traveler chooses an anticipating departure or not. In other words, if a traveler takes connection B with a completely reliable train service, and chooses connection A, for another unreliable set of services, then he has made an "anticipating departure". Similarly, the anticipating departure also exists when connection C is chosen under reliable services, and A or B is chosen under unreliable services.

7.3.2 Derivation of value of implied headway (marginal expected scheduling cost associated with headway)

Note that the scheduling disutility is generally not zero even for a completely reliable service. That is, whenever the train company does not provide the service that fits every traveler's PAT: for example, because PATs vary over travelers, there is always some disutility associated with scheduling inconvenience. Since we can compute the scheduling disutilities for the reliable train service under different levels of headway H, what is called the 'value of headway' can also be derived in our framework.

In Section 7.3.1, we obtain that the switching level of *PAT** for a choice between two reliable train connections is $\frac{\gamma \cdot H}{\beta+\gamma}$. A traveler chooses to take connection B if his *PAT* is earlier than *PAT**, and connection C if his PAT is later than PAT*. In such a case, *E[SDE]=PAT* and *E[SDL]=0*, when $0 \le PAT < \frac{\gamma \cdot H}{\beta+\gamma}$; *E[SDE]=0* and *E[SDL]=H-PAT* otherwise.

The expected scheduling costs for this reliable service can then be derived analytically for uniformly distributed PATs lying between 0 and H. To simplify the derivation, we again assume that $\theta = 0$. The derivation for the expected (average) scheduling costs, and its marginal value with respect to the headway, are shown in Eq.(7.7) and Eq.(7.8), respectively:

$$E[DU_s] = \frac{1}{H}\left[\int_0^{\frac{\gamma \cdot H}{\beta+\gamma}} \beta \cdot PAT \cdot dPAT + \int_{\frac{\gamma \cdot H}{\beta+\gamma}}^{H} \gamma \cdot (H - PAT) \cdot dPAT\right] = \frac{\beta \cdot \gamma}{2(\beta+\gamma)} H \tag{7.7}$$

$$\frac{\partial E[DU_s]}{\partial H} = \frac{\beta \cdot \gamma}{2(\beta+\gamma)}. \tag{7.8}$$

Note that (7.7) follows immediately as half the maximum schedule delay cost for the traveler at PAT*, which is $\beta \cdot PAT^* = \beta \cdot \gamma / (\beta + \gamma)$. The "half" stems from the linearity of schedule delay costs.

Eq.(7.7) is also given in Wardman (2004) (originally derived in Bates, 2003), where it is defined as the average headway disutility due to schedule inconvenience, and Eq.(7.8) can be regarded as the value of headway implied by schedule delay costs. It is clear that this headway value depends on the VSDE and VSDL, and it is intuitively plausible that higher values of schedule delay parameters would result in a higher value of headway (Eq.(7.8) is always positive).

7.3.3 Obtaining the value of implied schedule delay

In the previous subsection, we showed how to calculate the expected schedule delay costs for reliable services. We can calculate the expected schedule delay costs for unreliable train services in a similar way. By taking the difference of expected schedule delay costs between reliable and unreliable services, we are able to obtain the additional scheduling benefits (costs) associated with an improvement (or deterioration) of service reliability, and the value of changes in schedule delay costs resulting from the change in the standard deviation of travel times can also be obtained.

In determining the expected schedule delay costs for unreliable services, one crucial aspect in modeling is the travelers' departure connection choice. More specifically, if travelers are assumed to take the same connection for unreliable services as for reliable ones, the expected schedule delay costs will be different from the situation where travelers have anticipating departure, implying that travelers may change their departure behavior according to the levels of unreliability.

To illustrate the difference of computing the expected schedule delay costs in these two situations, we demonstrate two examples of the schedule delay cost functions for fixed and anticipating departures in Figures 7.2 and 7.3, respectively. In these two examples, headway is assumed to be 30 minutes, and the unreliable train is assumed to provide the service with a two mass-points travel time distribution with a 70 percent probability of arrival according to the schedule and a 30 percent probability of a 20-minute delay. For the schedule delay parameters, the coefficients obtained from Chapter 6 (see Table 6.5, scheduled commuting trips of Full model), with β_β = 4.68 (€/hour), β_γ = 28.56 (€/hour), and θ = 0, are used here. The black curves in Figures 7.2 and 7.3 depict the schedule delay costs (Eq.(7.1)) for the fully reliable services. These black curves also represent travelers' generalized travel costs (assumed as Eq.(7.9), the same utility specification as in Chapter 6 for scheduled commuting trips) if the travel time and cost are normalized to be zero. Note that PAT* determines the kink in this curve, because a traveler would choose to take connection B if his PAT is earlier than PAT*, and take connection C otherwise.

$$DU = \beta_C \cdot \mathrm{cost} + \beta_\alpha \cdot E[T] + \beta_\beta \cdot E[SDE] + \beta_\gamma \cdot E[SDL] + \beta_\sigma \cdot STD\,. \tag{7.9}$$

The gray curve (DU1) in Figure 7.2 represents the schedule delay cost for unreliable train services under the assumption that travelers' departure connection choice is the same as the one for reliable train services. This is what results when travelers ignore possible delays and just base their departure time on the official timetable. Since PAT* is the point where travelers are indifferent between choosing connections B and C for fully reliable services but not for unreliable services, there is a discrete jump of the cost function in the gray curve at PAT*. This discontinuity reflects that with, unreliable services, PAT* is not the optimal switching point between services B and C. Because there is a probability of delay, travelers with a PAT equal to PAT* are now still better-off with the earlier connection B, and therefore would face discretely higher expected scheduling costs when choosing C, whereas the gray line only reflects schedule delay costs. Travel delay costs also become relevant when introducing travel time uncertainty, and the same holds for the value of pure unreliability. The gray dashed line (DU2) shows the total generalized travel costs (Eq.(7.9), including the value of average travel delay and of pure unreliability, using the value estimated in Chapter 6 that β_α = 11.71 (€/hour), β_σ = 4.26 (€/hour) (see Table 6.5, scheduled commuting trips of the Full model). Note that these additional costs are independent of a traveler's PAT, and therefore shift the earlier gray curves upwards, by a constant amount.

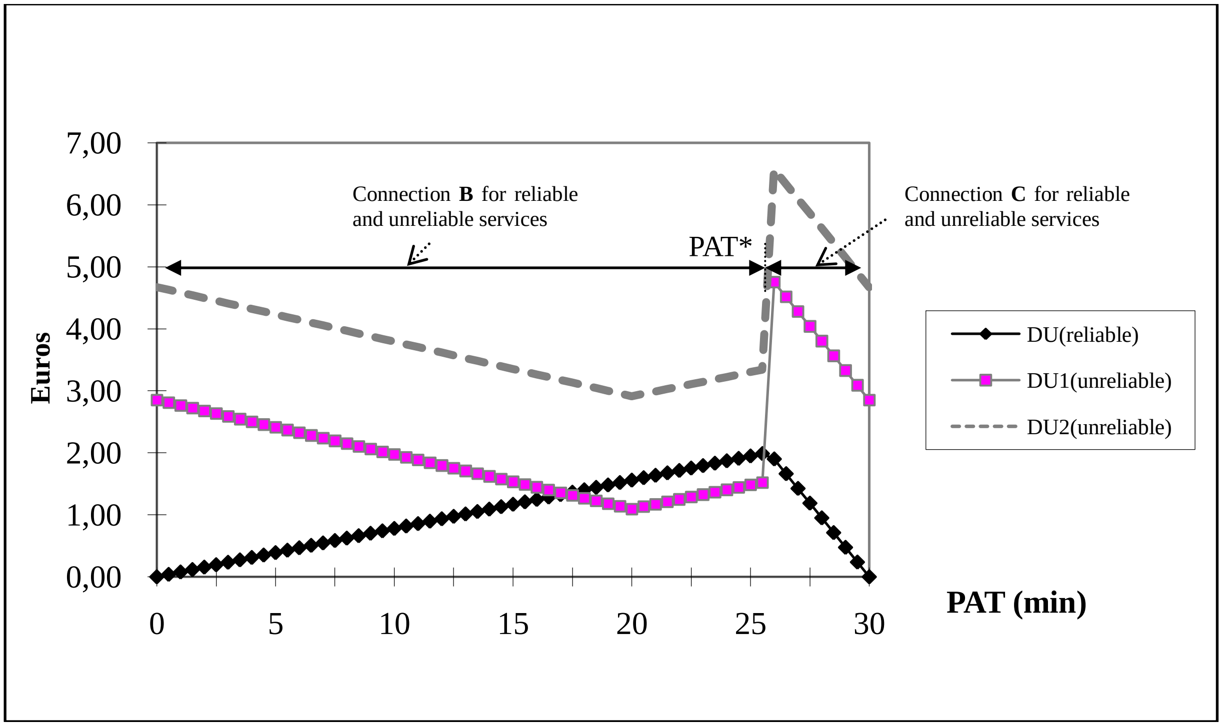


Figure 7.2 The schedule delay costs and generalized travel costs for fixed departure travelers

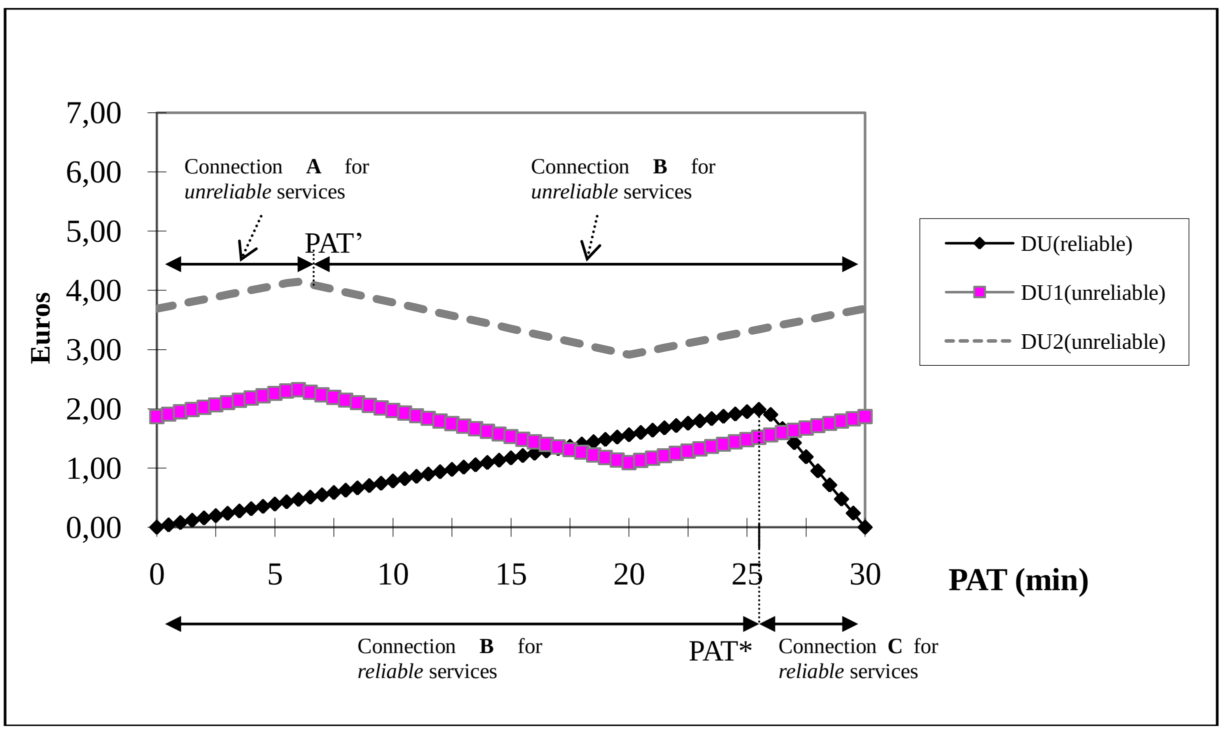


Figure 7.3 The schedule delay costs and generalized travel costs for anticipating departure travelers

Next, the gray curve in Figure 7.3 depicts the schedule delay costs for unreliable services when travelers' anticipating choice of departure time is taken into account. In such a case, since travelers will choose the connection that yields the smallest disutility according to the level of unreliability, the resulting connection choice may be different from the case when service is fully reliable. Just as for PAT* in the reliable service, it is also possible to determine the switching level of PAT for the unreliable service. PAT' in Figure 7.3 represents the point at which travelers are indifferent between choosing connections A and B in the unreliable services. In other words, a traveler would choose connection A when his PAT is earlier than PAT', otherwise, connection B is chosen. Note that the gray line is now continuous, indicating that the switch between connections is made at the optimal moment. Again, we also show the generalized travel cost level (gray dashed line, denoted as DU2) with the value of average travel delay and of pure unreliability included.

Once the schedule delay function has been derived for every relevant PAT, we can easily calculate the expected (average) schedule delay costs for all travelers under both reliable and unreliable situations. By taking the difference of expected schedule delay costs between these two services, the average additional schedule delay costs due to a particular level of unreliability can be obtained, and the value of implied schedule delay can be computed in terms of the associated standard deviation of the unreliability level, as shown in Eq.(7.10):

$$VOISD = \frac{E[DU_s]^U - E[DU_s]^R}{STD}, \tag{7.10}$$

where VOISD denotes the value of implied schedule delay, and $E[DU_s]^U$ and $E[DU_s]^R$ represent the expected schedule delay costs (disutilities) for unreliable and reliable services, respectively.

Since the value of implied schedule delay is defined as the marginal change in schedule delay costs associated with the change in the standard deviation of travel times, it provides a link for transferring Small's scheduling model to the mean-variance model. The utility specification in Eq.(7.9) can therefore be rewritten as Eq.(7.11):

$$DU = \beta_C \cdot \mathrm{cost} + \beta_\alpha \cdot E[T] + (\mathrm{VOISD} + \beta_\sigma) \cdot STD \tag{7.11}$$

7.4 NUMERICAL ILLUSTRATIONS

In order to analyze numerically anticipating departures, and to derive the value of implied scheduling costs, specific assumptions of travel times, travelers' PAT distribution, and estimates of the cost coefficients β, γ, θ are needed. We will use the simplest case of the two mass-points travel time distribution. Table 7.1 lists the range of various delay times and probabilities that we examine in this section, and the associated standard deviations are also given in this table. For the estimation of schedule delay parameters, the coefficients obtained from Chapter 6, with β = 4.68 (€/hour), γ = 28.53 (€/hour), and θ = 0, will be used for the numerical illustration. For the distribution of PATs, we assume that it is uniform. Travelers' preferred arrival times at the destination stations are independent of the scheduled timetable.

Table 7.1 The standard deviations for various levels of delay time and probability

Standard deviation						
Delay time	**Delay probability**					
	5%	**10%**	**20%**	**30%**	**40%**	**50%**
5 min	1.09	1.50	2.00	2.29	2.45	2.50
10 min	2.18	3.00	4.00	4.58	4.90	5.00
15 min	3.27	4.50	6.00	6.87	7.35	7.50
20 min	4.36	6.00	8.00	9.17	9.80	10.00
25 min	5.45	7.50	10.00	11.46	12.25	12.50
30 min	6.54	9.00	12.00	13.75	14.70	15.00

Table 7.2 summarizes the resulting percentages of travelers' anticipating departure under various travel time distributions (that is, given various pairs of delay time and delay probability). A plausible result is that the percentage of anticipating departures increases as the delay time and /or the delay probability increase. The effect of delay time and probability on the percentage of anticipating departures is nearly linear when the level of delay time and probability are moderate. However, the anticipating percentage increases much more than proportionally when the level of delay time and probability both become very large. In other words, travelers' anticipating departures become very common when the services are extremely unreliable.

Table 7.2 Percentage of travelers' anticipating departures for various levels of unreliability

Headway = 30 min, β = 4.68, γ = 28.53, and θ = 0						
Delay time	**Delay probability**					
	5%	**10%**	**20%**	**30%**	**40%**	**50%**
5 min	0.8%	1.7%	3.3%	5.0%	6.7%	8.3%
10 min	1.7%	3.3%	6.7%	10.0%	13.3%	19.3%
15 min	2.5%	5.0%	10.0%	17.2%	28.9%	35.9%
20 min	3.3%	6.7%	13.3%	33.8%	45.6%	52.6%
25 min	4.2%	8.3%	27.0%	51.0%	62.3%	69.3%
30 min	4.6%	9.6%	43.7%	67.2%	78.9%	85.9%

The value of headway implied by schedule delay costs can be calculated by using Eq.(7.8) in Section 7.3.3. Using the schedule delay parameters estimated from commuters' scheduled trips in Chapter 6 (as in Eq.(7.9)), β_β = 4.68 €/hour and β_γ = 28.53 €/hour, we get 2.10 €/hour for value of headway (VOH). In Chapter 6, we also estimated the VOH directly for non-scheduled trips (with disutility function $DU = \beta_c \cdot C + \beta_\alpha \cdot E[T] + \beta_\sigma \cdot STD + \beta_H \cdot H$). It is interesting to compare the derived and estimated VOHs. Table 7.3 gives the VOH estimated from Chapter 6 (Column II) and the VOH derived in this chapter (Column I). The last row shows the "headway valuation ratio", which is defined, analogously to the reliability ratio, as VOH/VOT. The headway valuation ratios in Table 7.3 reveal that the ratio derived from the schedule delay costs is slightly lower than the one estimated directly from the choice model. One of the reasons may be that these two sets of estimates were obtained from different types of trips by different groups of respondents, and thus the headway was evaluated differently. Another explanation could be that travelers favor a higher frequency of train services not only because it reduces the schedule delay costs but because it also helps travelers feel more flexible in planning their trips or reduces the risk of missing the train; consequently, the headway valuation ratio might be higher in the direct estimate than in the one derived solely from the scheduling consideration.

Table 7.3 The implied value of headway based on the estimates in Chapter 6 and (€/hour)

	Column I **Model for commuters scheduled trips**	**Column II** **Model for non-scheduled trips**
Value of time (VOT)	11.71	5.26
Value of schedule delay early	4.68	
Value of schedule delay late	28.53	
Value of headway (VOH)	**2.10**[a]	**1.51**[b]
Value of reliability (standard deviation)	4.26	11.64
VOH/VOT	**0.18**	**0.29**

Notes: [a.] Derived from the cost of schedule delay (see Eq.(7.8); [b.] Estimated directly from the choice model.

Next, under the same hypothesized reliability levels, headway, and schedule delay parameters, we can calculate the values of implied schedule delay as discussed in Section 7.3.3. We consider both the case where travelers are assumed to have fixed departures and the case where they have anticipating departures. The results are summarized in Table 7.4 for various levels of unreliability. The detailed derivation of the value of implied schedule delay for a two mass-points travel time distribution is provided in Appendix 7A. We then calculate the ratio between the value of implied schedule delay and the value of time, and refer to this ratio as the 'schedule delay valuation ratio' (SDR), which is similar to the definition of the reliability ratio (RR) (see Chapter 2). The resulting ratios for various levels of unreliability are given in Table 7.5.

In Tables 7.4 and 7.5 it is clear that the values of implied schedule delay or the schedule delay valuation ratios are not constant over different levels of unreliability. Different from the case of car journeys examined by Noland and Small (1995), who found that the value of implied schedule delay is independent of the standard deviation (under certain assumptions of travel time distributions), our results show that the value of implied schedule delay varies with the level of unreliability. This also confirms the conclusion in Fosgerau and Karlström (2007), who found that the expected schedule delay cost is not linear in the standard deviation of travel time for a scheduled service. In the case of fixed departures, the value of implied schedule delay increases as the unreliability level, either by delay time or delay probability, increases, and the value rises rapidly when the train service becomes extremely unreliable. As for the case of anticipating departures, the variation pattern is different from the one in the case of fixed departures: the resulting value of implied schedule delay does not increase monotonically with the level of unreliability. In general, this value increases with the level of unreliability when that level is moderate; nevertheless, the value slightly decreases with the delay probability when the probability is more than 30 percent in our numerical example. This may be because, when train services are extremely unreliable, most travelers just adjust their departure time to take an earlier connection; consequently, the expected schedule delay costs and the resulting value of schedule delay would decrease instead. Note that, when comparing the results between fixed and anticipating departures, the values overall are lower in the case of anticipating departures. This makes sense, as a traveler with anticipating departure is assumed to choose his departure connection optimally according to the reliability level so that his scheduling cost is minimized.

Table 7.4 Values of implied schedule delay for various levels of unreliability (€/hr)

Headway = 30 min, β = 4.68, γ = 28.53, and θ = 0						
Fixed departure						
Delay time	Delay probability					
	5%	10%	20%	30%	40%	50%
5 min	0.64	0.93	1.39	1.82	2.27	2.78
10 min	1.27	1.85	2.77	3.63	4.53	5.55
15 min	1.91	2.77	4.16	5.44	6.79	8.31
20 min	2.54	3.69	5.54	7.25	9.05	11.08
25 min	3.18	4.62	6.92	9.07	11.31	13.85
30 min	3.74	5.43	8.14	10.66	13.30	16.28
Anticipating departure						
Delay time	Delay probability					
	5%	10%	20%	30%	40%	50%
5 min	0.60	0.83	1.11	1.27	1.36	1.38
10 min	1.21	1.66	2.21	2.54	2.71	2.70
15 min	1.81	2.49	3.32	3.79	3.71	3.36
20 min	2.41	3.32	4.43	4.63	4.22	3.69
25 min	3.13	4.15	5.40	5.13	4.52	3.89
30 min	3.59	4.97	6.06	5.47	4.72	4.02

Table 7.5 Schedule delay valuation ratios for various levels of unreliability

Headway = 30 min, α =11.71, β = 4.68, γ = 28.53, and θ = 0						
Fixed departure						
Delay time	Delay probability					
	5%	10%	20%	30%	40%	50%
5 min	0.05	0.08	0.12	0.16	0.19	0.24
10 min	0.11	0.16	0.24	0.31	0.39	0.47
15 min	0.16	0.24	0.36	0.46	0.58	0.71
20 min	0.22	0.32	0.47	0.62	0.77	0.95
25 min	0.27	0.39	0.59	0.77	0.97	1.18
30 min	0.32	0.46	0.70	0.91	1.14	1.39
Anticipating departure						
Delay time	Delay probability					
	5%	10%	20%	30%	40%	50%
5 min	0.05	0.07	0.09	0.11	0.12	0.12
10 min	0.10	0.14	0.19	0.22	0.23	0.23
15 min	0.15	0.21	0.28	0.32	0.32	0.29
20 min	0.21	0.28	0.38	0.40	0.36	0.32
25 min	0.27	0.35	0.46	0.44	0.39	0.33
30 min	0.31	0.42	0.52	0.47	0.40	0.34

In Eq.(7.11) we showed that the scheduling model can be transformed into the mean-variance model by using the value of implied schedule delay. It is therefore possible to calculate the reliability ratio based on the transformed mean-variance model, as shown in Eq.(7.12). To distinguish our derived reliability ratio from the one obtained in the ‘true’ mean-variance model, we define the ratio in Eq.(7.12) as the 'generalized reliability valuation ratio' (GRR). Table 7.6 summarizes the generalized reliability ratios across various levels of reliability.

$$GRR = \frac{VOISD + \beta_\sigma}{\beta_\alpha}. \tag{7.12}$$

Recall that we have derived the implied RR from the meta-analysis in Chapter 2. It is therefore interesting to see the comparison between the result from this chapter and the one from the previous literature. In Table 2.9, the implied RR for the mean-variance model is around 1.71, which is generally higher than what we obtained in Table 7.6. Nevertheless, when we compare the values in Table 7.6 with the estimated mean-variance model result in Chapter 6 (see Table 6.5, with $\beta_\alpha = 16.62$ (€/hour) and $\beta_\sigma = 9.22$ (€/hour), and RR = 0.59)), we find that it is close to the generalized reliability valuation ratios we derived for anticipating departures, ranging from 0.42 to 0.88. There could be two different explanations. The first one is that the RR we estimated from Chapter 6 is indeed smaller than what was found in the literature, although there are only a few empirical VOR studies in the context of rail transport. The second possibility is that assuming all travelers have "anticipating departures" is rather conservative in the derivation of schedule delay costs, as the resulting schedule delay costs under the "fixed departures" assumption are higher than those in "anticipating departures". This explanation may be true when there are many inexperienced travelers in the population, since less experienced travelers would usually not be able to adjust their departure time according to the changes in the unreliability level of train services.

Table 7.6 Generalized reliability valuation ratios for various levels of unreliability

Headway = 30 min, α=11.71, β = 4.68, γ = 28.53, and θ = 0						
Fixed departure						
Delay time	Delay probability					
	5%	10%	20%	30%	40%	50%
5 min	0.42	0.44	0.48	0.52	0.56	0.60
10 min	0.47	0.52	0.60	0.67	0.75	0.84
15 min	0.53	0.60	0.72	0.83	0.94	1.07
20 min	0.58	0.68	0.84	0.98	1.14	1.31
25 min	0.64	0.76	0.95	1.14	1.33	1.55
30 min	0.68	0.83	1.06	1.27	1.50	1.75
Anticipating departure						
Delay time	Delay probability					
	5%	10%	20%	30%	40%	50%
5 min	0.42	0.43	0.46	0.47	0.48	0.48
10 min	0.47	0.51	0.55	0.58	0.60	0.59
15 min	0.52	0.58	0.65	0.69	0.68	0.65
20 min	0.57	0.65	0.74	0.76	0.72	0.68
25 min	0.63	0.72	0.83	0.80	0.75	0.70
30 min	0.67	0.79	0.88	0.83	0.77	0.71

7.5 APPLICATION IN COST-BENEFIT ANALYSIS

It is of interest to see what our analysis implies for cost-benefit analysis (CBA) of policies that reduce unreliability. CBA serves as a framework for evaluating the economic and social effects of large-scale investment in transport projects. While quantifying benefits from travel time savings has been a standard practice in the CBA framework, it is not entirely clear how to incorporate the benefit from the reliablity improvement. In previous chapters (Chapters 4 and 6) we showed that travelers do place significant values on reliablity (standard deviation), or schedule delay, or both. Thus, the lack of values of reliability would sometimes seriously limit the acceptance of the outcomes of the CBA and may raise doubts regarding the precision of the results.

The improvement of service reliability can result in the reduction of the standard deviation of travel time, as well as the reduction in (expected) schedule delay costs. In Sections 7.3 and 7.4, we showed how to derive the expected scheduling costs and to translate the schedule delay costs into the values based on the standard deviation definition. In this section we investigate the benefit contributions of each item with respect to the reliability improvement of travelers' generalized cost for a train trip.

An example is given in Table 7.7, in which there is a 30 percent probability of having a 20-minute delay for an unreliable train service with 30 minutes headway. By using the estimated parameters from commuters' scheduled trips in Chapter 6, we are able to calculate the benefits (costs) for the expected travel time, expected schedule delay, and standard deviation (pure reliability) reduction, respectively, if this unreliable service is improved to become a fully reliable one. Since the expected schedule delay costs are different between different departure pattern assumptions, i.e. fixed and anticipating departures as discussed in previous sections, it is of interest to see the impact of taking into account travelers' behavior change on the schedule delay benefit in relation to the total benefit. Table 7.7 shows the benefits for each item, and their related contribution to the total benefit of travelers' generalized costs.

Table 7.7 An example of the benefit contribution for reliability improvement[a]

	Fixed departure		**Anticipating departure**	
	Benefit (in euros)	**% of contribution**	**Benefit (in euros)**	**% of contribution**
Expected travel time reduction (estimated α = 11.71 €/hr)	1.171	**39.97%**	1.171	**46.30%**
Expected scheduling delay reduction (estimated β = 4.68 €/hr; γ = 28.53€/hr)	1.108	**37.82%**	0.707	**27.97%**
Standard deviation reduction (estimated σ = 4.26 €/hr)	0.651	**22.21%**	0.651	**25.73%**
Total benefit	2.930	100%	2.529	100%

[a.] Based on a delay probability of 30% and a delay time of 20 minutes.

In this example, the benefit contribution from the reliability improvement (sum of the benefits of the expected schedule delay and standard deviation reductions) is more than the contribution from the expected travel time reduction in both the fixed and the anticipating departures scenarios. Since the travel time savings is usually one of the largest components in the CBA assessment, these results suggest that the typical CBA procedure, which fails to take into account the benefits from reliability improvement, may seriously bias the outcome. The example in Table 7.7 shows that the underestimation is about 60 percent and 54 percent of total benefits for fixed and anticipating departure travelers, respectively. Note that the schedule delay benefit is smaller in the case of anticipating departure than it is in the case of fixed departure. This implies that, when travelers' anticipating departure is taken into account, the resulting benefit from the schedule delay reduction will not be as large as the situation in which travelers' departure behavior is assumed unchanged, and optimized for a fully-reliable service. This is intuitive: one can reverse the change and then the results show that an increase in unreliability without allowing optimal anticipation induces higher expected schedule delay costs than if this optimization is allowed.

To see how sensitively the reliability level influences the benefit contributions, we calculate the percentages of benefit contributions for a complete elimination of unreliability, starting at different initial levels of unreliability. Figure 7.4 demonstrates the variations of the benefit contributions when travelers' departure times are assumed fixed under both reliable and unreliable services. Figure 7.5 shows the same indicators when travelers' behavior changes, i.e. anticipating departures, are taken into account. The full sets of percentages of benefit contributions are provided in Tables 7B1 and 7B2 in Appendix 7B.

With fixed departures, as shown in Figure 7.4, it is obvious that the percentage of benefit contribution from schedule delay increases monotonically as the level of unreliability increases, whereas the percentage of benefit contribution from standard deviation (pure reliability) decreases with the unreliability level. The variations in the percentage of benefit contribution from the expected travel time range from 25 percent to 62 percent across various levels of unreliability. This suggests that, if policy practitioners only take into account the benefit from expected travel time reduction and ignore that from reliability improvement, the resulting total benefit from travelers' generalized cost may be underestimated substantially.

With anticipating departures, as shown in Figure 7.5, it should be noted that the variation of the percentage of benefit contribution from schedule delay is similar to the variation of value of implied schedule delay obtained in Table 7.4. The same reasoning discussed in the previous section also applies here – when train services are extremely unreliable, travelers are more inclined to exhibit anticipating departure and, consequently, the expected schedule delay cost would be smaller than it would be with fixed departures. In this case, the percentage contribution from the expected travel time ranges from 25 percent to 67 percent across various levels of unreliability. Again, similar to the results for fixed departures, ignoring the benefit from reliability improvement will seriously bias the CBA result. When we compare the expected travel time contribution between the cases of fixed and anticipating departures, the difference only becomes noticeable when the unreliability level is large. This stems from the impact of anticipating departure behavior on the expected schedule delay costs when services are very unreliable. On the other hand, this also implies that, when the degree of unreliability is moderate, failing to take travelers' anticipating departures, and changes therein, into account will not lead to much difference in the result of the benefit calculation. But ignoring scheduled delay costs and costs of pure unreliability remains, of course, a potentially large source of underestimation of benefit contribution.

Finally, our result in this section is comparable to what was found in Ettema and Timmermans (2006), who concluded that scheduling costs account for 20 to 40 percent of the generalized costs. As shown in Tables 7B1 and 7B2, given a wide range of unreliability levels, the schedule delay costs contribute 8 to 50 percent and 8 to 38 percent to the total benefits in the case of fixed and anticipating departures, respectively. It seems that our finding here is not far from the result of the analysis in Ettema and Timmermans (2006), even though these two studies are in the different contexts (train versus car transport) and are calculated from different schedule delay parameters settings.

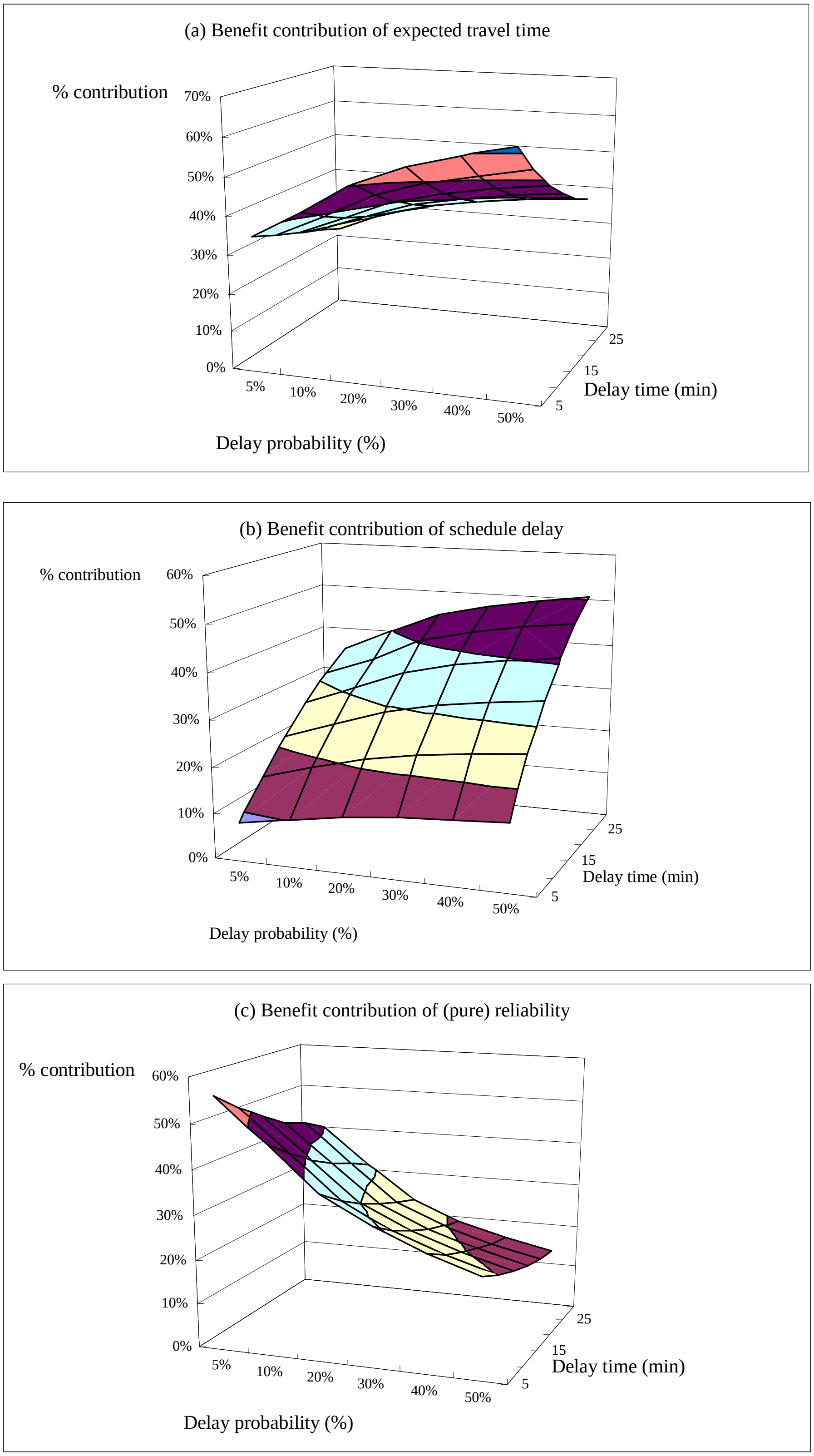


Figure 7.4 Graphical illustration of the benefit contributions for (a) expected travel time, (b) schedule delay, and (c) (pure) reliability reductions in the case of fixed departure

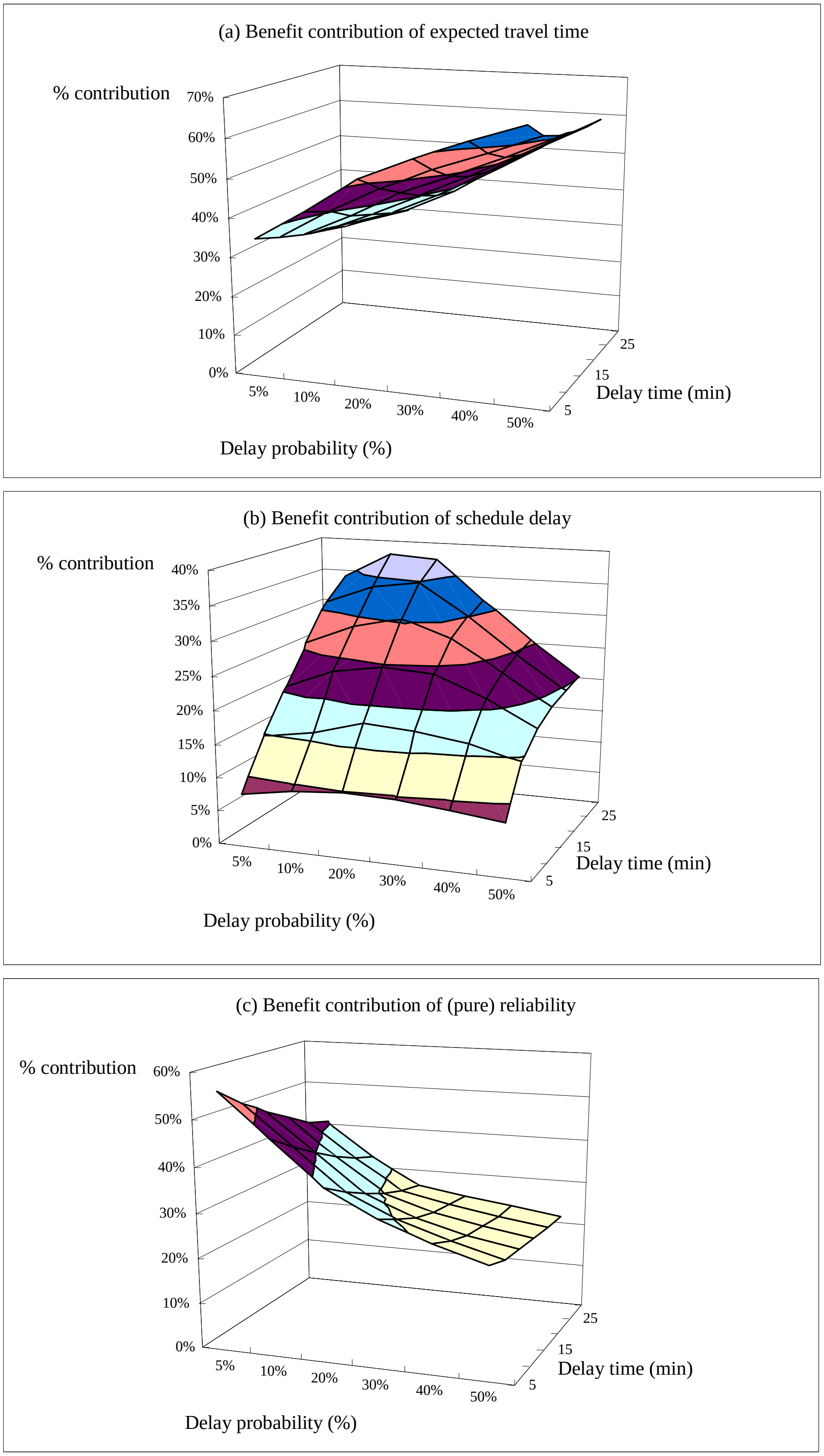


Figure 7.5 Graphical illustration of the benefit contributions for (a) expected travel time, (b) schedule delay, and (c) (pure) reliability reductions in the case of anticipating departure

7.6 CONCLUSIONS

We established a simple model to describe and predict railway passengers' anticipating departure time behavior given their preferred arrival times, when choosing between scheduled train connection arrival times. Furthermore, we derived the expected schedule delay costs with and without considering this anticipating departures behavior for some specific (two mass-points) travel time distributions. The values of implied schedule delay can be related to the standard deviation of travel time distribution. The numerical examples in Section 7.4 provide some insight into how the anticipating departures and the values of implied schedule delay vary with the levels of unreliability.

In reality, public transport operators may be more interested in some travel time distributions than in others; our analysis can of course be applied to any multi-mass-points distribution so that a value can be determined for any specific application. The result can also be generalized by trying many possible (and reasonable) levels of unreliability, headways, and even different parameters of β, γ, and θ, so that a good approximation of an *average* value of implied schedule delay can be obtained statistically. The analysis in the final section focused on practical relevance. By exploring separate attributes in the benefit contributions of reliability improvement, it was possible to identify the amount of bias incurred with a naïve CBA that ignores the reliability benefit. For our parameterization, the underestimations range from 33 percent to 75 percent of total time-related benefits across various unreliability levels when travelers' anticipating departure behavior is taken into account.

Appendix 7A

Derivation for the value of implied schedule delay when the departure time is assumed to be fixed for both reliable and unreliable services

As discussed in Section 7.3.1, a traveler will choose only connections B or C in the case of fully reliable services, and the switching level of $PAT^* = \dfrac{\gamma \cdot H}{\beta + \gamma}$ (given that θ=0). It means that connection B is taken if a traveler's PAT falls in the interval $[0, \dfrac{\gamma \cdot H}{\beta + \gamma}]$, and connection C is taken otherwise. For the sake of simplicity, we will assume that θ=0 in the derivation below.

The expected scheduling cost for reliable service is (as shown Eq.(7.7)):

$$E[DU_s]^R = \frac{1}{H}\left[\int_0^{\frac{\gamma\cdot H}{\beta+\gamma}} \beta \cdot PAT \cdot dPAT + \int_{\frac{\gamma\cdot H}{\beta+\gamma}}^{H} \gamma \cdot (H - PAT) \cdot dPAT\right] = \frac{\beta \cdot \gamma}{2(\beta+\gamma)} H .$$

Suppose the traveler's departure time is fixed, regardless of the level of service reliability. In such a case, the switching point of PAT* is the same for both reliable and unreliable services. The expected scheduling costs for unreliable service can be derived separately under two situations:

a) When $DT < PAT^*$, the expected scheduling cost of unreliable service is

$$E[DU_s]^U = \frac{1}{H}\left[\begin{array}{l} \displaystyle\int_0^{DT} \left[\beta \cdot (1-p) \cdot PAT + \gamma \cdot p \cdot (DT - PAT)\right] \cdot dPAT \\ \displaystyle + \int_{DT}^{PAT^*} \left[\beta \cdot (1-p) \cdot PAT + \beta \cdot p \cdot (PAT - DT)\right] \cdot dPAT \\ \displaystyle + \int_{PAT^*}^{H} \gamma \cdot \left(H - PAT + p \cdot DT\right) \cdot dPAT \end{array}\right].$$

Thus, the difference of scheduling cost between unreliable and reliable train services is:

$$\Delta E[DU_s] = E[DU_s]^U - E[DU_s]^R = \frac{(\beta+\gamma)}{2H} \cdot p \cdot DT^2 .$$

b) When $DT \geq PAT^*$, the expected scheduling costs of unreliable service is:

$$E[DU_s]^U = \frac{1}{H}\left[\begin{array}{l} \int_0^{PAT*}\left[\beta\cdot(1-p)\cdot PAT+\gamma\cdot p\cdot(DT-PAT)\right]\cdot dPAT \\ +\int_{PAT*}^{H}\gamma\cdot\left(H-PAT+p\cdot DT\right)\cdot dPAT \end{array}\right].$$

Thus, the difference in scheduling costs between unreliable and reliable train services is:

$$\Delta E[DU_s] = E[DU_s]^U - E[DU_s]^R = \gamma\cdot p\cdot DT - \frac{\gamma^2}{2(\beta+\gamma)}\cdot p\cdot H\ .$$

Since the standard deviation of two mass-points distribution is:

$$STD = DT\cdot\sqrt{(1-p)\cdot p^2 + p\cdot(1-p)^2}\ ,$$

we can easily obtain the value of implied scheduling costs (i.e. the marginal expected scheduling cost associated with standard deviation) as $\frac{\Delta E[DU_s]}{STD}$.

Derivation of the value of implied schedule delay when the anticipating departure is taken into account for the unreliable service

When the traveler's behavior change, i.e. anticipating departure, is taken into account, the switching level of PAT for unreliable services is different from the case for reliable services. Furthermore, connection A becomes one of the available choices among the alternative connections in the unreliable services, such that there are two switching levels of PAT in this case. Here, we define *PAT'* as the switching point between A and B, and *PAT''* as the switching point between B and C.

The switching level of PAT between connections A and B is:

$$PAT'' = DT - \frac{\beta\cdot H}{(\beta+\gamma)\cdot p}\ .$$

The switching level of PAT between connections B and C is:

$$PAT' = \frac{\gamma\cdot H}{\beta+\gamma} + p\cdot DT\ .$$

For this unreliable train service, there is a critical level of delay time, denoted as *DT'*, such that, when $DT < DT^*$, connection A is always worse-off than connection B, whilst when $DT \geqq DT^*$, connection C is always worse-off than connection B. This critical delay time can be obtained as:

$$DT^* = \frac{\beta \cdot H}{(\beta+\gamma)\cdot p}.$$

Thus, when $DT < DT^*$, the traveler will only choose between connections B and C, and, when $DT \geqq DT^*$, the traveler will therefore only choose between connections A and B. The expected scheduling costs for this unreliable service can therefore be computed separately for these two situations:

a) When $DT < DT^*$, connections B and C will be chosen, the expected scheduling cost for unreliable service is:

$$E[DU_s]^U = \frac{1}{H}\left[\begin{array}{l} \int_0^{DT}\left[\beta\cdot(1-p)\cdot PAT - \gamma\cdot p\cdot(DT-PAT)\right]\cdot dPAT \\ + \int_{DT}^{PAT'}\left[\beta\cdot(1-p)\cdot PAT + \beta\cdot p\cdot(PAT-DT)\right]\cdot dPAT \\ + \int_{PAT'}^{H}\gamma\cdot\left(H-PAT+p\cdot DT\right)\cdot dPAT \end{array}\right].$$

The difference in scheduling cost between unreliable and reliable train services is:

$$\Delta E[DU_s] = E[DU_s]^U - E[DU_s]^R = \frac{(\beta+\gamma)}{2H}\cdot p\cdot(1-p)\cdot DT^2.$$

b) When $DT \geqq DT^*$, connections A and B will be chosen, and the expected scheduling cost for unreliable service is:

$$E[DU_s]^U = \frac{1}{H}\left[\begin{array}{l} \int_0^{PAT''}\beta\cdot(PAT+H-p\cdot DT)\cdot dPAT \\ + \int_{PAT''}^{DT}\left[\beta\cdot(1-p)\cdot PAT + \gamma\cdot p\cdot(DT-PAT)\right]\cdot dPAT \\ + \int_{DT}^{H}\left[\beta\cdot(1-p)\cdot PAT + \beta\cdot p\cdot(PAT-DT)\right]\cdot dPAT \end{array}\right].$$

The difference in scheduling costs between unreliable and reliable train services is:

$$\Delta E[DU_s] = E[DU_s]^U - E[DU_s]^R = \beta \cdot \left(\frac{1}{2} H + (1-p) \cdot DT - \frac{\beta \cdot H}{2(\beta+\gamma) \cdot p} \right) - \frac{\beta \cdot \gamma \cdot H}{2(\beta+\gamma)} .$$

Again, we can easily obtain the value of implied scheduling costs (i.e. the marginal expected scheduling cost associated with the standard deviation) as $\frac{\Delta E[DU_s]}{STD}$.

Appendix 7B

Table 7B1 Percentages of benefit contribution in the case of fixed departure

Benefit of average travel time reduction						
Delay time	*Delay probability*					
	5%	*10%*	*20%*	*30%*	*40%*	*50%*
5 min	35.42%	42.93%	50.90%	55.77%	59.42%	62.46%
10 min	32.68%	38.99%	45.43%	49.28%	52.10%	54.42%
15 min	30.34%	35.70%	41.02%	44.14%	46.39%	48.22%
20 min	28.32%	32.92%	37.40%	39.97%	41.81%	43.26%
25 min	25.54%	30.93%	34.36%	36.52%	38.05%	39.27%
30 min	25.15%	28.72%	32.07%	33.94%	35.26%	36.30%
Benefit of schedule delay						
Delay time	*Delay probability*					
	5%	*10%*	*20%*	*30%*	*40%*	*50%*
5 min	8.42%	10.21%	12.07%	13.24%	14.10%	22.72%
10 min	15.48%	18.46%	21.52%	23.34%	24.68%	25.78%
15 min	21.54%	25.34%	29.13%	31.33%	32.94%	34.24%
20 min	26.79%	31.15%	35.39%	37.82%	39.56%	41.00%
25 min	31.38%	35.31%	40.64%	43.19%	45.00%	46.44%
30 min	34.97%	39.94%	44.60%	47.20%	49.03%	50.49%
Benefit of (pure) reliability						
Delay time	*Delay probability*					
	5%	*10%*	*20%*	*30%*	*40%*	*50%*
5 min	56.16%	46.86%	37.03%	30.99%	26.48%	14.82%
10 min	51.84%	42.55%	33.05%	27.38%	23.22%	19.80%
15 min	48.12%	38.96%	29.85%	24.53%	20.67%	17.54%
20 min	44.89%	35.93%	27.21%	22.21%	18.63%	15.74%
25 min	43.08%	33.76%	25.00%	20.29%	16.95%	14.29%
30 min	39.88%	31.34%	23.33%	18.86%	15.71%	13.21%

Table 7B2 Percentages of benefit contribution in the case of anticipating departure

Benefit of average travel time reduction						
Delay time	*Delay probability*					
	5%	*10%*	*20%*	*30%*	*40%*	*50%*
5 min	35.58%	43.40%	52.18%	58.10%	63.00%	67.48%
10 min	32.95%	40.64%	47.49%	53.00%	57.83%	62.71%
15 min	30.68%	36.64%	43.57%	48.78%	54.53%	60.57%
20 min	28.70%	33.99%	40.26%	46.30%	53.00%	59.56%
25 min	26.68%	31.70%	37.73%	44.94%	52.13%	58.97%
30 min	25.50%	29.72%	36.20%	44.07%	51.56%	58.58%
Benefit of schedule delay						
Delay time	*Delay probability*					
	5%	*10%*	*20%*	*30%*	*40%*	*50%*
5 min	7.99%	9.23%	9.86%	9.61%	8.93%	7.97%
10 min	14.80%	16.00%	17.96%	17.54%	16.40%	14.47%
15 min	20.67%	23.38%	24.72%	24.12%	21.17%	17.39%
20 min	25.78%	28.92%	30.45%	27.97%	23.38%	18.77%
25 min	31.03%	33.71%	34.82%	30.09%	24.64%	19.58%
30 min	34.05%	37.85%	37.47%	31.44%	25.47%	20.11%
Benefit of (pure) reliability						
Delay time	*Delay probability*					
	5%	*10%*	*20%*	*30%*	*40%*	*50%*
5 min	56.43%	47.37%	37.96%	32.29%	28.07%	24.55%
10 min	52.25%	43.36%	34.55%	29.46%	25.77%	22.82%
15 min	48.65%	39.98%	31.71%	27.10%	24.30%	22.04%
20 min	45.52%	37.09%	29.29%	25.73%	23.62%	21.67%
25 min	42.29%	34.59%	27.45%	24.97%	23.23%	21.45%
30 min	40.45%	32.43%	26.33%	24.49%	22.97%	21.31%

CHAPTER 8

8 CONCLUSION

8.1 SUMMARY

The unreliability of travel time is regarded as an important factor in travelers' decision making on their mode choice, route choice, and departure time choice. Travel time unreliability may generate considerable generalized costs for a traveler. The valuation of travel time reliability is the central topic of this thesis. Empirical estimates of travel time reliability are found to vary widely in the literature, and no consensus seems to have been achieved either on the measures of reliability to use in applied work or on the modeling approaches.

This thesis started with an attempt to explain the variation in findings through meta-analytical procedures (see *Chapter 2*). Here we address Research Question 1 identified in the introduction, which concerns the factors influencing estimates of reliability ratio (RR), schedule delay early ratio (SDER), and schedule delay late ratio (SDLR). (These are estimates of the values divided by the value of time (VOT).) The most important factor is the utility specification. The meta-regressions strongly suggest that including both reliability and schedule variables would have a significant negative effect on RR estimates, as well as on SDER and SDLR estimates. This result confirms expectations, since the concepts of reliability and schedule delay are closely related, and statistically, they are highly correlated. Thus, if researchers mis-specify individuals' utility functions, the possible bias of the estimates could be quite substantial.

The mode and trip characteristics also have some impacts on the schedule delay ratios. The commuting trip is usually considered to be one with relatively strong scheduling concerns. Our results concerning the SDLR confirm this. Regarding the effect of travel modes, we find that there is a significant negative impact on the scheduling ratio estimates for the public transport mode. One possible explanation is that public transport users usually have to accommodate the preferred arrival times to the presented timetable. As such, these people have less restricted preferred arrival times than the private car users, so the penalty of being early or late is relatively small for public transport than it is for car. Another finding is that the RR estimated by Min-Max (uncertain travel time bounds) measurement is considerably higher than the one estimated by other measurements. One possible explanation could be that the travelers interpret the 'Min-Max' and 'unreliability' (some possible travel time realizations) attributes differently in the choice experiment than what is intended by the researchers.

Next, this thesis investigates how respondents perceive and interpret travel time reliability, and tests their understanding of different ways of presenting reliability (see *Chapter 3*, which addresses Research Question 2) by means of in-depth interviews. In the analyses eight different formats were tested: some use bar charts, some use clock-face presentations, and some only use verbal descriptions. The interviews focused not only on respondents' assessments of these formats regarding clarity, ease of handling, and visual attractiveness, but also on their understanding of the reliability information presented in these formats. The results suggest that the format with verbal descriptions is preferred and best understood by the majority of respondents. Moreover, it is also equally favored by lower- and higher-educated people. This format is therefore recommended for the future large-scale valuation of reliability survey in the Netherlands.

The central question in this thesis concerns the empirical estimates of reliability and schedule delay (see *Chapter 4* and *Chapter 6*, which both empirically address Research Question 3a and 3b). *Chapter 4* analyzes the choice behavior among Dutch morning car commuters under a road-pricing scheme. The main stated choice experiment (SCE) consisted of 11 choice sets, and respondents were asked to allocate 10 trips over 4 alternatives – 3 car and 1 public transport alternatives – where the levels of the attributes were based on the current behavior of each individual. Various choice models were estimated by using the choice proportions setup. When we include scheduling costs in the estimations, "pure uncertainty" becomes insignificant. This suggests that the effects of reliability can be fully explained by scheduling considerations in the context of road transport, specifically for car commuters. The nonlinear effects of scheduling variables were also addressed in our model estimations. The analysis indicates that people's aversion to arriving early increases nonlinearly as their schedule delay early time increases.

The resulting parameter values for VOT (8.9 €/hour) and VSDL (14.6 €/hour) seem rather plausible, though the VSDE (12.0 €/hour) estimate is surprisingly high. This may be explained by the nonlinear effect of the SDE variable. The VOT, VSDE, and VSDL of public transport are significantly different from those of car transport. The VOT is lower for car, whereas the VSDE and VSDL are significantly higher. Various covariates were also studied. Most of the findings are in line with the earlier literature, such as a positive trip length effect on the VOT; positive income effects on the VOT; and positive gender (female) effects on the VSDE and VSDL. In particular, the schedule delay costs are lower for respondents with a higher income and a higher educational level. The explanation for this may be that higher-educated people have higher-status jobs, which generally have less restricted working times. Travel cost compensation only seems to have an impact on the VOT, with fully-compensated commuters generally having higher VOT.

Chapter 6 analyzed the choice behavior of Dutch railway passengers. A tailored choice experiment was constructed, based on the individual's current travel experience, so that the tradeoff in the choice situation was meaningful to the respondent. In the estimation of scheduled trips, the utility specification that includes both reliability and schedule delay variables is our most-preferred model. When comparing the estimates between commuters and non-commuters, we find that commuters' VOT, VSDE, and VSDL are generally higher than those for non-commuters. This suggests that time and scheduling are more costly for commuting trips than for other trip purposes. Regarding the analysis of covariate effects, income has a positive effect on the VOT and VOR, although sometimes the effect is insignificant. Education, as expected, has a similar effect to income. Female railway users have a lower VOT and VOR. Trip length again has a significant positive impact on the VOT and VOR. For scheduled commuting trips, the VOT, VSDE, VSDL, and VOR increase considerably with the levels of cost compensation.

In *Chapter 5* an alternative, dynamic framework for estimating time-varying values of travel time savings and values of schedule delay was proposed. Research Question 4, which concerns the impact of different utility specifications on the values of time and reliability, is addressed here. The results suggest that individuals' time-related shadow prices vary strongly over the morning peak, and values of travel time savings are consequently strongly time-dependent. A failure to incorporate such considerations may produce biased estimates of values of travel time savings, and errors in the prediction of behavioral responses to policies or other measures that affect the time pattern of congestion in the morning peak.

Finally, *Chapter 7* addressed Research Question 5a and 5b, which concerns the train travelers' anticipating departure behavior, and the application of the schedule delay parameters in practice. In this chapter, a simple model is established to describe and predict railway passengers' anticipating departure time behavior when train services are unreliable. The expected schedule delay costs and the value of implied schedule delay were derived analytically, with and without considering this anticipating departures behavior, for some specific (two-mass points) travel time distributions. Based on the parameterization in Chapter 6, the resulting reliability ratio (RR) derived from this analytical framework is comparable to the one obtained directly in Chapter 6. The final section of Chapter 7 focused on the practical relevance of the findings. By exploring separate attributes in the benefit contributions of reliability improvement, the resulting benefit underestimation from the traditional CBA, which ignores the reliability and/or scheduling benefits, can be up to 50 percent or more, depending on the initial levels of unreliability and the parameters. Although this figure is derived for a specific situation, it demonstrates that the possible bias incurred by ignoring the reliability and/or schedule delay benefits can be quite substantial.

8.2 RELEVANCE TO THE RESEARCH AND POLICY

This thesis aimed first to make a number of contributions to the research in this area. First, the meta-analysis on the existing empirical VOR studies contributes to a deeper understanding of the effects of different reliability measures and modeling approaches on the estimates of the reliability ratio and schedule delay ratios. Second, the thesis investigates how travelers perceive and interpret travel time reliability in their trip making, and how they understand the reliability information presented in the SP, by means of in-depth interviews. Third, the empirical part of this thesis explores the impacts of different modeling approaches and various covariates on the variations of the reliability and schedule delay estimates, both for road and rail transport. Fourth, the extension of Small's trip scheduling model to account for time-varying values of time provides an alternative modeling framework for estimating the values of time and schedule delay, and the resulting time-varying shadow prices of time-related variables suggests that strong time-preferences do exist for the morning commuters across the peak hours, which cannot be properly captured with the conventional linear model. Lastly, the application of the schedule delay estimates in modeling railway passengers' anticipating departure behavior is, as far as we are aware, unique.

From a policy perspective, this thesis contributes to the implementation of including reliability and scheduling benefits into CBAs. Since reliability is becoming a central focus in transportation policy in many countries, it is essential that the benefits gained from reliability improvements can be incorporated into project evaluation frameworks. The analytical model presented in Chapter 7 gives the opportunity to derive such benefits. Incorporating the benefits from reliability improvements can result in a more accurate evaluation for transportation investments. As such, the transport options provided by the policy makers can be more cost-effective and be better matched with the users' preferences. For example, investment in rapid response incident clearance systems, which mainly improves the reliability in travel times but not the average travel times, may prove more cost-effective and attractive than the alternative of expanding highway capacity projects.

The second policy-relevant aspect of the research relates to the implementation of traffic assignment procedures, and this can be done, for example, by incorporating the effects of travel time reliability into route choice models. Travelers' responses to the changes in reliability levels can therefore be better predicted; at the same time, the practitioners may provide a better means for reducing congestion or improving the network capacity utilization.

The final practical relevance of the research concerns the determination of optimal road prices. The value of reliability and/or schedule delay would play a role in determining the optimal toll of road pricing, as different levels of road use result in different levels of unreliability, and thus encompass a congestion-type externality.

8.3 FUTURE RESEARCH RECOMMENDATIONS

We started this thesis by remarking that there are a number of unresolved issues in the knowledge of travel time reliability valuation. We have addressed and analyzed some of these issues, as discussed in Section 8.1, and have contributed to a better understanding of how to assess travel time reliability in the context of personal travel. Nevertheless, various interesting research questions in this area remain to be answered. In this section, we will mention some of these briefly.

The first issue concerns the choice data type used for valuing travel time reliability. In determining the VOR, stated preference (SP) approaches appear a logical choice, as also recommended in Bates et al. (2001). The main reason is that it is difficult to collect travel time reliability information for chosen and non-chosen alternatives for revealed preference (RP) data. Moreover, travel time, reliability, and cost are usually highly correlated, which hampers the estimation for separate parameters or model convergence. Still, RP data is based on travelers' real behavior and actual mode/route shares. Some recent US studies, e.g. Small et al. (2005), where RP data were successfully collected from the choices between tolled (variably) non-tolled congested routes, have shown the advantages of combining both SP and RP data in estimating the VOT and VOR. Good quality RP data can help to validate SP results, and vice versa. Estimating a model based on the pooled RP and SP data may help to exploit the strengths and to ameliorate the weaknesses of these two data sources. Thus, it would be ideal, once road pricing is implemented in the Netherlands, to collect RP data through automated vehicle identification, in order to accurately track travelers' actual behaviors, as well as travel time reliability. These travelers could then also be approached to participate in an SP questionnaire.

A second interesting research question relates to the modeling approach, i.e. utility specification, of individual's choice behavior. The analyses in Chapters 4, 5, and 6 show that the empirical estimates of travel time, reliability and schedule delay variables depends very much on how the utility function is specified. From the literature survey in Chapter 2, it is evident that there is still no general consensus on the best representative measure for reliability and on the best way of modeling the effects of reliability. Certainly, the appropriate measure and utility specification may strongly depend on the formulation of the choice questions and data, and thus varies across different studies. Because of the complex nature of the concept of reliability, it is always challenging for researchers to make the choice questions as comprehensible as possible, but at the same time keep sufficient information in the choice questions that can be used flexibly to test a great numbers of utility specifications. To answer this research question will require much broader and much more extensive empirical findings in this area.

Another interesting research direction is the use of schedule delay costs in practical policy evaluation. The concept of schedule delay has been well recognized as a good way to describe the consequences of unreliability, and its costs have to be included in the utility specification. Although we have demonstrated how to derive the value implied by schedule delay in Chapter 7, this is, however, done only for some specific cases. Hence, future research is needed to apply the schedule delay estimates in a more general and broadly accepted way.

Our final recommendation for future research on reliability valuation concerns freight transport. Compared with the VOR research in passenger transport, the modeling approaches and the measures used to describe unreliability are more diverse in freight transport. In addition, it is clear that travel time unreliability affects both shippers and carriers. Thus, the risk-sharing between shippers and carriers may play a role in reliability valuation. All of this requires much more empirical research.

REFERENCES

Amelsfort, D. van, 2004, Description of choice experiment, internal working document, Delft.

Abdel-Aty, M., R. Kitamura, and P. Jovanis, 1995, Investigating the effect of travel time variability on route choice using repeated measurement stated preference data, *Transportation Research Record* 1493, pp. 39–45.

Arnott, R., A. de Palma, and R. Lindsey, 1993, A structural model of peak-period congestion: a traffic bottleneck with elastic demand, *American Economic Review* 83 (1) 161-179.

Arup Transport Planning, 2002, Modelling and appraisal of journey time variability, *Report for UK Department for Transport (DfT).*

Ashiru, O., J.W. Polak, and R.B. Noland, 2004, The utility of schedules: A theoretical model of departure time choice and activity time allocation with application to individual activity schedules, *Transportation Research Record* 1894 pp. 84-98.

Avineri, E., and J. N. Prashker, 2005, Sensitivity to travel time variability: Travelers' learning perspective, *Transportation Research Part C 13*, pp. 157-183.

AVV (Adviesdienst Verkeer en Vervoer) 2005, The Value of Reliability in Transport: Provisional values for the Netherlands based on expert opinion, Rotterdam.

AVV (Adviesdienst Verkeer en Vervoer) 2006, Value of Time Personenvervoer en Value of Reliability Personenvervoer,http://www.rws-avv.nl/servlet/page?_pageid=116&_dad=portal30&_schema=PORTAL30&p_item_type=theme&p_theme_name=PERSONENVERVOER, accessed 9 June 2006.

Bates, J.J., 1987, Measuring travel time values with a discrete choice model: a note, *Economic Journal* 97, pp.493-498.

Bates, J., 2003, Departure Time Shifts, Waiting and Headway, Appendix B in Faber Maunsell (2003).

Bates, J., J. Fearon, and I. Black, 2003, Frameworks for modeling the variability of journey times on the highway network: a report for UK DfT, December 2003.

Bates J., J. Polak, P. Jones, and A. Cook, 2001, The valuation of reliability for personal travel, *Transportation Research Part E* 37, pp. 191-229.

Batley, R. 2007, Marginal valuations of travel time and scheduling, and the reliability premium, *Transportation Research Part E* 43, pp. 387-408.

Becker, G.S., 1965, A theory of the allocation of time, *Economic Journal* 75, pp.493-517.

Ben-Akiva, M. and S. R. Lerman, 1997, *Discrete choice analysis*, The MIT Press, Cambridge.

Bhat, C. R. and Sardesai R., 2006, The impact of stop-making and travel time reliability on commute mode choice, *Transportation Research Part B* 40, pp. 709-730.

Black, I.G. and J.G. Towriss 1993, Demand effects of travel time reliability, Center for Logistics and Transportation, Cranfield Institute of Technology

Brownstone, D. and K. A. Small, 2003, Valuing time and reliability: assessing the evidence from road pricing demonstrations, University of California, Irvine, Working paper.

Chang, X.B. and P. R. Stopher, 1981, Defining the perceived attributes of travel modes for urban work trips, *Transportation Planning and Technology* 7, pp. 55–65.

Cascetta, E. and A. Papola, 2003, A joint mode-transit service choice model incorporating the effect of regional transport service timetables, *Transportation Research Part B* 37, pp. 595-614.

Chin, A. T. H., 1990, Influences on commuter trip departure time decisions in Singapore, *Transportation Research A* 24 (5), pp. 321-333.

Cooper, H. and L.V. Hedges, 1994, *The Handbook of Research Synthesis*, New York: Russel Sage Function.

Copley, G., P. Murphy, and D. Pearce, 2002, Understanding and valuing journey time variability. *Proceedings of the European Transport Conference* – 2002, Cambridge, PTRC, London.

DeSerpa, A.C., 1971, A theory of the economics of time, *Economic Journal* 81, pp. 828-846.

Ettema, D., G. Tamminga, H. Timmermans, and T. Arentze, 2005, A micro-simulation model system of departure time using a perception updating model under travel time uncertainty, *Transportation Research Part A* 39, pp. 325-344.

Ettema, D. and H.J.P. Timmermans, 2003, Modeling departure time choice in the context of activity scheduling behaviour, *Transportation Research Record* 1831, pp. 39-46.

Ettema, D., and H. Timmermans, 2006, Costs of travel time uncertainty and benefits of travel time information: Conceptual model and numerical examples, *Transportation Research Part C* 14, pp. 335-350.

Florax, R.J.G.M., C. M. Travisi, and P. Nijkamp, 2005, A meta-analysis of the willingness to pay for reductions in pesticide risk exposure, *European Review of Agricultural Economics* 32(4), pp. 441-467.

Fosgerau, M. and A. Karlström, 2007, The value of reliability, MPRA paper No. 5733, Online at http://mpra.ub.uni-muenchen.de/5733/.

Gaver, Jr. D.P., 1968, Headstart strategies for combating congestion, *Transportation Science* 2, pp. 172-181.

Ghosh, A., 2001, Valuing time and reliability: commuters' mode choice from a real time congestion pricing experiment, PhD Dissertation, University of California, Irvine.

Greene, W. H., 2003, *Econometric Analysis*, fifth edition, Pearson Education, Inc.

Gunn, H., 2001, Spatial and temporal transferability of relationships between travel demand, trip cost and travel time, *Transportation Research E* 37 (18), pp. 163-198.

Hamer, R.N., G. C. de Jong, E. P. Kroes, P. Warffemius, ,2005, The value of reliability in transport. Outcomes of an expert workshop, *RAND Europe Report TR-240-AVV*, Leiden.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper and L. V. Hedges (eds.), The Handbook of Research Synthesis. New York: Russell Sage Foundation, pp. 285–299.

Hendrickson, C. and E. Plank, 1984, The flexibility of departure times for work trips, *Transport Research Part A* 18 (1), pp. 25-36.

Hensher, D. A., 2001, The valuation of commuter travel time savings for car drivers: evaluating alternative model specification, *Transportation* 28, pp. 101-118.

Hensher, D.A., and W. H. Greene, 2003, The mixed logit model: the state of practice, *Transportation* 30 (2), pp. 133-176.

Hensher, D. A., J. M. Rose, and W. H. Greene, 2005, *Applied choice analysis: A primer*, Cambridge University Press, Cambridge.

Hess, S., J. W. Polak, A. Daly, and G. Hyman, 2007, Flexible substitution patterns in models of mode and time of day choice: new evidence from the UK and the Netherlands, *Transportation* 34, No.2, pp. 213-238.

Hollander, Y., 2005, The Attitudes of Bus Users to Travel Time Variability, Proceedings of the European Transport Conference – 2005, Strasbourg. PTRC, London.

Hollander, Y., 2006, Direct versus indirect models for the effects of unreliability, *Transportation Research Part A* 40, pp. 699-711.

Jackson, W. B. and J. V. Jucker, 1981, An empirical study of travel time variability and travel choice behavior, *Transportation Science* 16 (4), pp. 460-475.

Jiang, M. and T. Morikawa, 2004, Theoretical analysis on the variation of value of travel time savings, *Transportation Research A* 38 (8), pp.551-571.

Johnson, M., 1966, Travel time and the price of leisure, *Western Economic Journal*, Spring, pp. 135-145.

Jong, G. de, A. Daly, M. Pieters, C. Vellay, M. Bradley, and F. Hofman, 2003, A model for time of day and mode choice using error components logit, *Transportation Research Part E* 39, pp. 245-268.

Kahneman, D. and A. Tversky, 1979, Prospect theory: an analysis of decision under risk, *Econometrica* 47, pp. 263-291.

Kahneman, D. and A. Tversky, 2000, *Choices*, *Values and Frames*, Cambridge University Press, Cambridge.

Knight, T. E., 1974, An approach to the evaluation of changes in travel time unreliability: a "safety margin" hypothesis, *Transportation* 3 (4), pp. 393-408.

König, A., and K. W. Axhausen, 2002, The reliability of the transportation system and its influence on the choice behaviour, Swiss Transport Research Conference (STRC), Monte Verità/Ascona.

Koskenoja, P. M. K., 1996, The effect of unreliable commuting time on commuter preferences, PhD Dissertation, Department of Economics, University of California, Irvine.

Lam, T.C., K. A. Small, 2001, The value of time and reliability: measurement from a value pricing experiment, *Transportation Research Part E* 37, pp. 231-251.

Lint, J. W. C. van and H. J. van Zuylen, 2005, Monitoring and predicting freeway travel time reliability, *Transportation Research Record* 1917, pp. 54-62.

Lint, J. W. C. van, H. J. van Zuylen, and H. Tu, 2008, Travel time unreliability on freeways: Why measures based on variance tell only half the story, *Transportation Research Part A* 42, pp. 258-277.

Liu, H. X., W. Recker, and A. Chen, 2004, Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data, *Transportation Research Part A* 38, pp. 435-453.

Liu, H. X., X. He, and W. Recker, 2007, Estimation of time-dependency of values of travel time and its reliability from loop detector data, *Transportation Research* 41B, pp. 448-461.

Louviere, J. J., D. A. Hensher, and J. D. Swait, 2000, *Stated choice methods*, Cambridge University Press, Cambridge.

McFadden, D., and K. Train, 2000, Mixed MNL models for discrete response, *Journal of Applied Econometrics* 15, pp. 447-470.

Noland, R. B., and J. W. Polak, 2002, Travel time variability: a review of theoretical and empirical issues, *Transport Review* 22 (1), pp. 39-54.

Noland, R. B. and K. A. Small, 1995, Travel time uncertainty, departure time and the cost of the morning commute, Paper presented to the 74th annual meeting of the *Transportation Research Board*, Washington, Jan. 1995.

Noland, R. B., K. A. Small, P. M. Koskenoja, X. Chu, 1998, Simulating travel reliability, *Regional Science and Urban Economics* 28 (5), pp. 535-564.

Oort, C.J., 1969, The evaluation of travelling time, *Journal of Transport Economics and Policy* 3, pp. 279-286.

Palma, A. de, C. Fontan, and N. Picard, 2003, Departure time choice: estimates results and simulation to Paris area, 10th International Conference on Travel Behavior Research, Lucern, August 2003

Palma, A. de, and R. Lindsey, 2001, Optimal timetables for public transportation, *Transportation Research Part B* 35, pp. 789-813.

Polak, J., 1987, Travel time variability and departure time choice: a utility theoretic approach, Transport Studies Group, Discussion paper No. 15, Polytechnic of London.

Prashker, J.N., 1979, Direct analysis of the perceived importance of attributes of reliability of travel modes in urban travel, *Transportation* 8, pp. 329–346.

RAND Europe, 2004, De Waardering van Kwaliteit en Betrouwbaarheid in personen – en goederen vervoer [The valuation of quality and reliability in passenger and freight transport], AVV/RAND Europe, Rotterdam.

Revelt, D., and K. Train, 1998, Mixed logit with repeated choices: households' choices of appliance efficiency level, *The Review of Economics and Statistics* 80 (4), pp. 647-657.

Rietveld, P., F. R. Bruinsma, and D. J. van Vuuren, 2001, Coping with unreliability in public transport chains: a case study for Netherlands, *Transportation Research Part A* 35, pp. 539-559.

Redmond, L. S. and P. L. Mokhtarian, 2001, The positive utility of the commute: modeling ideal commute time and relative desired commute amount, *Transportation* 28, No.2, pp. 179-205.

Robbins L., 1932, An essay on the nature and significance of economic science, London: Macmillan.

Senbil, M. and R. Kitumura, 2004, Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames, *Intelligent Transportation Systems* 8, pp. 19-31.

Senna, L. A. D. S., 1994, The Influence of Travel Time Variability on the Value of Time, *Transportation* 21, pp.203-228.

Shires, J. D. and G. C. de Jong, 2008, An international meta-analysis of values of travel time savings, Institute for Transportation Studies, University of Leeds.

Small, K. A., 1982, The scheduling of consumer activities: work trips, *American Economic Review* 72, pp. 467-479.

Small, K. A., R. B. Noland, and P. Koskenoja, 1995, Socio-economic attributes and impacts of travel reliability: a stated preference approach, University of California, Irvince, California PATH research report.

Small, K. A., R. B. Noland, X. Chu, and D. Lewis., 1999, Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation, *Transportation Research Board NCHRP Report* 431.

Small, K. A., C. Winston, and J. Yan, 2002, Uncovering the distribution of motorists' preferences for travel time and reliability: implications for road pricing, University of California Transportation Center, working paper #546.

Small, K. A., C. Winston, J. Yan, 2005, Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability, *Econometrica* 73 (4) , 1367–1382

Small, K.A. and E. T. Verhoef, 2007, *The Economics of Urban Transportation*, Routledge, London.

Stanley, T.D., 2001, Wheat from Chaff: Meta-Analysis as Quantitative Literature Review, *The Journal of Economic Perspectives* 15 (3), pp. 131-150.

Tilahun, N., and D. Levinson, 2005, Travel Time Reliability in Route Choice using Adaptive Stated Preference; working paper, Department of Civil Engineering, University of Minnesota.

Tseng, Y.-Y., B. Ubbels, and E. T. Verhoef, 2005, Value of time, schedule delay, and reliability – Estimation results of a stated choice experiment among Dutch commuters facing congestion. Paper presented at the 45th Congress of European Regional Science Association.

Tseng, Y.-Y., E. T. Verhoef, and P. Rietveld, 2006, Valuation of travel time reliability for railway passengers ─ A stated choice experiment for Dutch railway travelers, Paper presented at the *seminar Infrastructure Reliability TU Delft* in June 2006, and *NECTAR conference* in Finland in July 2006.

Ubbels, B., 2006, Road pricing: effectiveness, acceptance and institutional aspects, PhD dissertation, Department of Spatial Economics, Vrije Universiteit Amsterdam, the Netherlands.

Vickrey, W.S., 1969, Congestion theory and transport investment, *American Economic Review (Papers and Proceedings)* 59, pp. 251-260.

Von Neumann, J. and O. Morgenstern, 1944, *Theory of Games and Economics Behavior*, Princeton University Press, Princeton.

Wardman, M., 2001, A review of British evidence on time and service quality valuations, *Transportation Research Part E* 37, pp. 107-128.

Wardman, M., 2004, Public transport values of time, *Transport Policy* 11, pp. 363-377.

Wilson, P. W., 1989, Scheduling costs and the value of travel time, *Urban Studies* 26, pp. 356-366.

Yan, J., 2002, Heterogeneity in motorists' preferences for travel time and time reliability: empirical finding from multiple survey data sets and its policy implications, PhD Dissertation, Department of Economics, University of California, Irvine.

SAMENVATTING (Summary in Dutch)

De onbetrouwbaarheid van reistijden word gezien als een belangrijke factor in het besluitvormingsproces van reizigers met betrekking tot hun vervoerswijze, routekeuze en keuze van het vertrektijdstip. Reistijdonbetrouwbaarheid veroorzaakt aanzienlijke gegeneraliseerde kosten voor de reiziger. De waardering van reistijdonbetrouwbaarheid vormt het centrale thema van deze thesis. Uit de literatuur blijkt dat empirische schattingen van reistijdonbetrouwbaarheid erg verschillen en dat er nog geen consensus is met betrekking tot de te hanteren maatstaf van onbetrouwbaarheid in toegepast onderzoek of modelstudies. Deze studie begint met een poging om de variatie in empirische schattingen te verklaren met behulp van meta-analytische methodes. Hierbij onderzoeken we de rol van factoren die de schattingen van de onbetrouwbaarheids ratio (RR), de ‘schedule delay early’ ratio (SDER), de ‘schedule delay late’ ratio (SDLR) beïnvloeden. (Het betreft hier schattingen van de waarden gedeeld door de waarde van tijd (VOT).) De belangrijkste invloedsfactor is de specificatie van de nutsfunctie. De resultaten van de metaregressie-analyses laten duidelijk zien dat wanneer zowel onbetrouwbaarheids- als tijdstipskeuze-variabelen worden toegevoegd, de schattingen van zowel RR, SDER als SDLR significant lager uitvallen. Dit resultaat sluit aan bij de verwachtingen aangezien onbetrouwbaarheid en tijdstipkeuze conceptueel sterk gerelateerd, en statistisch gezien gecorreleerd zijn. Als onderzoekers de nutsfunctie verkeerd specificeren kan de mogelijke fout in de schattingen daarom groot zijn. Ook de karakteristieken van de vervoerswijze en de routekarakteristieken zijn van invloed op de ‘schedule delay’-ratio’s. Woon-werkverplaatsingen worden normaal gesproken beschouwd als verplaatsingen met relatief strikte beperkingen ten aanzien van de aankomsttijd. Onze resultaten aangaande de SDLR bevestigen dit. Voor wat betreft de invloed van vervoerswijze vinden we een significant negatieve invloed op de geschatte ‘schedule delay’ ratio’s voor het openbaar vervoer. Een mogelijke verklaring hiervoor is dat gebruikers van het openbaar vervoer hun gewenste aankomsttijden aan moeten passen aan de dienstregeling. Daardoor hebben deze mensen minder beperkingen voor hun aankomsttijden dan reizigers die met de auto gaan, waardoor de penalty van een te vroege of te late aankomst relatief gering is. Een ander resultaat is dat de geschatte RR hoger is wanneer deze wordt geschat op basis van de Min-Max-methode (onzekere reistijdgrenzen) dan op basis van andere meetmethoden. Een mogelijke verklaring is dat reizigers bij het keuze-experiment de attributen ‘Min-Max’ en ‘onbetrouwbaarheid’ (een aantal mogelijke reistijdrealisaties) anders interpreteren dan de onderzoeker dit bedoeld heeft.

Vervolgens onderzoeken we op basis van diepte-interviews hoe respondenten reistijdonbetrouwbaarheid beleven en interpreteren en testen we hun begrip van verschillende manieren om reistijdonbetrouwbaarheid weer te geven. In de analyse zijn acht verschillende formats getest; sommigen op basis van staafdiagrammen, anderen op basis van ‘clock-face’ representaties en weer anderen op basis van alleen een verbale beschrijving. De interviews richten zich niet alleen op de mening van de respondent over de duidelijkheid van de formats, het gebruiksgemak en de visuele aantrekkelijkheid, maar ook op het begrip van de informatie over de onbetrouwbaarheid. Het blijkt dat respondenten de verbale beschrijving het meest waarderen en het best begrijpen. Tevens wordt dit format zowel door laag- als hoogopgeleide mensen begrepen. Er wordt dan ook aangeraden dit format te gebruiken voor toekomstig grootschalig onderzoek naar de waardering van onbetrouwbaarheid in Nederland.

De centrale vraag in deze studie betreft de empirische schatting van onbetrouwbaarheid- en kosten van tijdstipkeuze (zie hoofdstuk 4 en hoofdstuk 6). Hoofdstuk 4 analyseert het keuzegedrag van Nederlandse ochtendspitsreizigers die moeten betalen voor het weggebruik. Het belangrijkste stated choice-experiment (SCE) bestaat uit 11 keuzesituaties. Aan respondenten werd gevraagd om 10 reizen te verdelen over 4 alternatieven, waarvan 3 betrekking hebben op de auto en 1 openbaar vervoer. De waarde van de attributen waren hierbij gebaseerd op het huidige gedrag van het betreffende individu. Verschillende keuzemodellen werden geschat met behulp van de ‘choice proportions’-opzet. Als we de kosten van tijdstipkeuze in de schattingen opnamen werd de ’pure onzekerheid’ insignificant. Dit suggereert dat de effecten van onbetrouwbaarheid volledig verklaard kunnen worden door tijdstipkeuze-overwegingen in de context van wegtransport, met name waar het woon-werkverkeer betreft. De niet-lineaire effecten van de tijdstipkeuzevariabelen zijn ook in het model opgenomen. De analyse laat zien dat de aversie om te vroeg te arriveren op niet-lineaire wijze toeneemt wanneer de ‘schedule delay early’ groter wordt. De geschatte waarden voor de VOT (8.9 €/uur) en VSDL (14.6 €/uur) lijken aannemelijk, hoewel de geschatte waarde voor de VSDE (12.0 €/uur) verrassend hoog is. Dit kan worden verklaard door het niet-lineaire effect van de SDE variabele. De VOT, VSDE en VSDL voor openbaar vervoer verschillen significant van de geschatte waarden voor autogebruikers. De VOT is lager voor de auto, terwijl de VSDE en VSDL significant hoger zijn. Ook hebben we de invloed van verschillende covariaten bestudeerd. De meeste bevindingen zijn vergelijkbaar met eerdere resultaten in de literatuur. De lengte van de reis heeft bijvoorbeeld een positief effect op de VOT, evenals het inkomensniveau. Verder geldt dat voor vrouwelijke reizigers de VSDE en de VSDL hoger zijn. Meer in het bijzonder geldt dat de kosten van tijdstipkeuze lager zijn voor respondenten met een hoger inkomen en een hoger opleidingsniveau. De verklaring hiervoor kan zijn dat deze mensen een betere baan hebben met vaak minder stringente werktijden. Reistijdcompensatie lijkt alleen effect te hebben op de VOT, waarbij geldt dat volledig gecompenseerde reizigers over het algemeen een hogere VOT hebben.

In hoofdstuk 6 analyseren we het keuzegedrag van de Nederlandse treinreiziger. Daarvoor gebruiken we een keuze-experiment ‘op maat’, gebaseerd op het huidige reisgedrag van de betreffende respondent zodat deze een betekenisvolle keuze kan maken. Voor schattingen waarbij tijdstipkeuze van belang is, kent het door ons geprefereerde model een nutsspecificatie waarin zowel betrouwbaarheids- als tijdstipvariabelen afzonderlijk zijn opgenomen.

Als we de schattingen voor forenzen met niet-forenzen vergelijken blijkt dat de VOT, VSDE en VSDL voor forenzen hoger is. Dit resultaat suggereert dat de kosten van reistijd- en ‘schedule delay’ over het algemeen hoger zijn bij woon-werkverplaatsingen dan bij andere reismotieven.

Uit de analyse van de invloed van covariaten blijkt dat het inkomensniveau een positief effect heeft op de VOT en de VOR, hoewel het effect soms insignificant is. Opleidingsniveau heeft, zoals verwacht, een vergelijkbaar effect als het inkomensniveau. Vrouwelijke treinreizigers hebben een lagere VOT en VOR. De lengte van de reis heeft een significant positief effect op de VOT en de VOR. Voor woon-werkverplaatsingen waarbij tijdstipkeuze van belang is stijgen de VOT, VSDE, VSDL en VOR aanzienlijk wanneer het niveau van de kostencompensatie stijgt.

In hoofdstuk 5 introduceren we een alternatief dynamisch modelkader voor het schatten van tijdsafhankelijke waarden van reistijdbesparing en de waarde van ‘schedule delay’. De analyse in dit hoofdstuk betreft de impact van verschillende specificaties van de nutsfunctie op de waarden van tijd en reistijdonbetrouwbaarheid. De resultaten wijzen erop dat tijdsgerelateerde schaduwprijzen van individuen sterk variëren tijdens de ochtendspits en dat de waarde van reistijdbesparingen sterk tijdsafhankelijk is. Wanneer deze aandachtspunten niet opgenomen worden in het schattingsmodel, kan dit tot systematische schattingsfouten met betrekking tot de waarde van reistijdbesparingen leiden, alsmede tot foutieve voorspellingen van veranderingen in reizigersgedrag als gevolg van beleidsmaatregelen of andere maatregelen die het tijdspatroon van filevorming in de ochtendspits beïnvloeden.

Hoofdstuk 7, tenslotte, gaat over het anticiperend gedrag met betrekking tot vertrektijden door treinreizigers en de toepassing van de tijdstipkeuze-parameters in de praktijk. In dit hoofdstuk wordt een eenvoudig model ontwikkeld om het anticiperend gedrag door treinreizigers te beschrijven en te voorspellen wanneer de dienstregeling onbetrouwbaar is. De verwachte SDE-kosten en de waarde van ‘schedule delay’ worden voor een aantal specifieke reistijdverdelingen analytisch afgeleid, zowel met als zonder rekening te houden met anticiperend gedrag. Op basis van de parameterisering die in hoofdstuk 6 is gebruikt, is de analytisch berekende reliability ratio (RR) vergelijkbaar met de direct verkregen RR van hoofdstuk 6. Het laatste deel van hoofdstuk 7 richt zich op de praktische relevantie van de resultaten. Door verschillende attributen van de baten van betrouwbaarheidsverbeteringen te analyseren blijkt dat de onderschatting van traditionele kostenbatenanalyses, die geen rekening houden met baten van betrouwbaarheid en tijdstipkeuze, kan oplopen tot 50 procent of meer, afhankelijk van het beginniveau van de onbetrouwbaarheid en de parameters. Hoewel dit getal is afgeleid voor een specifieke situatie, laat dit resultaat zien dat de potentiële fout als gevolg van het negeren van de baten van betrouwbaarheid en tijstipkeuze aanzienlijk is.

RELEVANTIE VOOR ONDERZOEK EN BELEID

Deze thesis draagt op verschillende wijzen bij aan het onderzoek op dit gebied. Ten eerste draagt de meta-studie naar bestaande empirische VOR studies bij aan een beter begrip van de effecten van verschillende betrouwbaarheidsmaatstaven en modelbenaderingen op de reliability ratio (RR) en de ‘schedule delay’ ratio. Daarnaast wordt met behulp van diepte-interviews onderzocht hoe reizigers reistijdonbetrouwbaarheid beleven en interpreteren gedurende hun reis en in hoeverre ze de informatie van een SP-experiment begrijpen. Ten derde wordt in het empirische deel van deze studie onderzocht welk effect het gebruik van verschillende modelleringsmethodes en covariaten heeft op schattingen van de VOT en VOR voor zowel weg- als treinvervoer. Verder biedt de uitbreiding van Small's tijdstipkeuze model, waarbij rekening wordt gehouden met het feit dat de waardering van tijd per tijdstip kan varieren, een alternatieve methode om de waarde van tijd en 'schedule delay' te berekenen. De tijdsafhankelijke schaduwprijzen van de tijdsgerelateerde variabelen laten zien dat er voor ochtendspits reizigers sterke tijdspreferenties bestaan die met een conventioneel lineair model niet op een juiste manier gemodelleerd kunnen worden. Tenslotte is de toepassing van de 'schedule delay'- schattingen bij de modellering van anticiperend gedrag voor zover wij weten uniek.

Vanuit een beleidsperspectief draagt deze studie bij aan de implementatie van het opnemen van de baten van reistijdbetrouwbaarheid en tijdstipkeuze in kostenbatenanalyses. Omdat betrouwbaarheid in veel landen een centraal aandachtspunt binnen het vervoersbeleid is geworden, is het essentiëel dat de baten van betrouwbaarheidsverbetering in projectevaluaties kunnen worden opgenomen. Het analytische model in hoofdstuk 7 biedt de mogelijkheid om deze baten af te leiden. Het opnemen van de baten van betrouwbaarheid kan resulteren in een nauwkeurigere evaluatie van investeringen in de vervoerssector. Dit stelt beleidsmakers in staat voor kosteneffectievere vervoersmaatregelen te zorgen die beter aansluiten bij de wensen van de gebruikers. Investeringen in bijvoorbeeld de snelheid waarmee na een incident het spoor weer vrij gemaakt wordt, hetgeen met name de betrouwbaarheid verbetert maar niet de gemiddelde reistijd, kunnen meer kosteneffectief blijken dan het uitbreiden van de capaciteit van snelwegen.

Het tweede beleidsrelevante aspect van dit onderzoek betreft implicaties voor netwerkmodellering (in het bijzonder 'traffic assignment'). Dit kan bijvoorbeeld gedaan worden door de effecten van reistijdonbetrouwbaarheid in routekeuzemodellen te implementeren. Hiermee kan de reactie van reizigers op veranderingen in reistijdonbetrouwbaarheid beter voorspeld worden. Ook kunnen beleidsmakers betere manieren verzinnen om files terug te dringen en het netwerk beter te benutten.

Het laatste beleidsrelevante aspect van dit onderzoek betreft het bepalen van de optimale beprijzing van weggebruik. De waarden van onbetrouwbaarheid en/of tijdstipkeuze zullen een rol spelen in het bepalen van de optimale tol, aangezien verschillende niveaus van weggebruik resulteren in verschillende niveaus van onbetrouwbaarheid, hetgeen resulteert in een externaliteit zoals die bij fileproblematiek.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

389. G. DEBREZION, *Railway impacts on real estate prices.*
390. A.V. HARDIYANTO, *Time series studies on Indonesian rupiah/USD rate 1995 – 2005.*
391. M.I.S.H. MUNANDAR, *Essays on economic integration.*
392. K.G. BERDEN, *On technology, uncertainty and economic growth.*
393. G. VAN DE KUILEN, *The economic measurement of psychological risk attitudes.*
394. E.A. MOOI, *Inter-organizational cooperation, conflict, and change.*
395. A. LLENA NOZAL, *On the dynamics of health, work and socioeconomic status.*
396. P.D.E. DINDO, *Bounded rationality and heterogeneity in economic dynamic models*.
397. D.F. SCHRAGER, *Essays on asset liability modeling*.
398. R. HUANG, *Three essays on the effects of banking regulations*.
399. C.M. VAN MOURIK, G*lobalisation and the role of financial accounting information in Japan.*
400. S.M.S.N. MAXIMIANO, *Essays in organizational economics.*
401. W. JANSSENS, *Social capital and cooperation: An impact evaluation of a women's empowerment programme in rural India.*
402. J. VAN DER SLUIS, *Successful entrepreneurship and human capital.*
403. S. DOMINGUEZ MARTINEZ, *Decision making with asymmetric information*.
404. H. SUNARTO, *Understanding the role of bank relationships, relationship marketing, and organizational learning in the performance of people's credit bank.*
405. M.Â. DOS REIS PORTELA, *Four essays on education, growth and labour economics.*
406. S.S. FICCO, *Essays on imperfect information-processing in economics.*
407. P.J.P.M. VERSIJP, *Advances in the use of stochastic dominance in asset pricing.*
408. M.R. WILDENBEEST, *Consumer search and oligopolistic pricing: A theoretical and empirical inquiry.*

409. E. GUSTAFSSON-WRIGHT, *Baring the threads: Social capital, vulnerability and the well-being of children in Guatemala*.
410. S. YERGOU-WORKU, *Marriage markets and fertility in South Africa with comparisons to Britain and Sweden.*
411. J.F. SLIJKERMAN, *Financial stability in the EU*.
412. W.A. VAN DEN BERG, *Private equity acquisitions*.
413. Y. CHENG, *Selected topics on nonparametric conditional quantiles and risk theory*.
414. M. DE POOTER, *Modeling and forecasting stock return volatility and the term structure of interest rates*.
415. F. RAVAZZOLO, *Forecasting financial time series using model averaging.*
416. M.J.E. KABKI, *Transnationalism, local development and social security: the functioning of support networks in rural Ghana.*
417. M. POPLAWSKI RIBEIRO, *Fiscal policy under rules and restrictions.*
418. S.W. BISSESSUR, *Earnings, quality and earnings management: the role of accounting accruals.*
419. L. RATNOVSKI, *A Random Walk Down the Lombard Street: Essays on Banking.*
420. R.P. NICOLAI, *Maintenance models for systems subject to measurable deterioration.*
421. R.K. ANDADARI, *Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).*
422. V.KARTSEVA, *Designing Controls for Network Organizations: A Value-Based Approach.*
423. J. ARTS, Essays *on New Product Adoption and Diffusion.*
424. A. BABUS, *Essays on Networks: Theory and Applications.*
425. M. VAN DER VOORT, *Modelling Credit Derivatives.*
426. G. GARITA, *Financial Market Liberalization and Economic Growth.*
427. E.BEKKERS, Essays *on Firm Heterogeneity and Quality in International Trade.*
428. H.LEAHU, *Measure-Valued Differentiation for Finite Products of Measures: Theory and Applications.*
429. G. BALTUSSEN, *New Insights into Behavioral Finance.*
430. W. VERMEULEN, *Essays on Housing Supply, Land Use Regulation and Regional Labour Markets.*
431. I.S. BUHAI, *Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions.*
432. C. ZHOU, *On Extreme Value Statistics.*
433. M. VAN DER WEL, *Riskfree Rate Dynamics: Information, Trading, and State Space Modeling.*

434. S.M.W. PHLIPPEN, *Come Close and Co-Create: Proximities in pharmaceutical innovation networks.*
435. A.V.P.B. MONTEIRO, *The Dynamics of Corporate Credit Risk: An Intensity-based Econometric Analysis.*
436. S.T. TRAUTMANN, *Uncertainty in Individual and Social Decisions: Theory and Experiments.*
437. R. LORD, *Efficient pricing algorithms for exotic derivatives.*
438. R.P. WOLTHOFF, *Essays on Simultaneous Search Equilibrium.*